``The central insight of categorical mathematics,"

Final draft submitted, for the volume *Invited talks of the Thirteenth International Congress of Logic, Methodology and Philosophy of Science, Beijing, 2007*, forthcoming.

May 29, 2008

THE CENTRAL INSIGHT OF CATEGORICAL MATHEMATICS

COLIN MCLARTY

Category theory grew from the long-standing use of *maps* or *morphisms* between mathematical structures. As an influential example, Richard Dedekind unified and extended earlier number theory by greater use of mappings in algebra. But he also defined the natural numbers themselves in purely structural terms by properties of the successor map. He asserted in foundations what he and others saw in practice: Neither practice nor foundations need say what the natural numbers are. Both must say how the natural numbers map to themselves and other sets.

Mathematics today uses *categories* of structures and morphisms. The general 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. Categories are structures in turn with *functors* as morphisms between them. Standard definitions characterize each structure not uniquely but *up to isomorphism* by its place in some pattern of maps. In research, structures are studied by these patterns of maps. Foundations can be given entirely in terms of these patterns. In ontology, mathematical objects need have no properties beyond the structural relations revealed by the maps. Category theory unified these as one insight: "It's the arrows that really matter!" (Awodey, 2006, p. 8).

Section 1 uses Riemann's complex analysis and other historical examples to motivate the basic ideas of morphism-based mathematics. This leads to a foundation in section 2 such that:

(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.

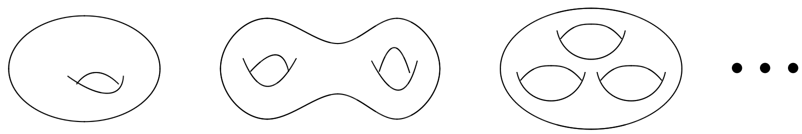
These first order axioms in the language of category theory suffice as a *formal foundation* to define all the concepts and prove all the results of current mathematics. Besides that, though, no matter what formal foundation one prefers, the category of sets and the category of categories deserve study since they pervade the *working foundations* of current mathematics. For example section 2.1.1 defines a set of sets as a function the way geometers define a space of spaces as a map.

Theorem schemes in Section 3 show exactly what is meant by "structural" properties. Section 4 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).

1. Riemann surfaces and maps

Bernhard Riemann made huge advances in calculus using maps between simple spaces now called *Riemann Surfaces*. Indeed Riemann (1851) never says what the points of any surface are. He says how parts of each surface are mapped to the complex plane. He used two versions of maps: the broad class of all *continuous maps* and its very restrictive subclass of *holomorphic maps*.

Topology takes 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.

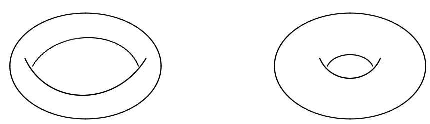


The number of holes is the *genus*, and surfaces are *topologically isomorphic* if and only if they have the same genus. Visually this means that no matter how two surfaces S_a and S_b are shaped, and how their holes are placed, they can stretch and bend to fit each other if and only they have equally many holes. Precisely: if they have the same genus then S_a maps onto S_b by some continuous $f\colon S_a \to S_b$ while S_b maps onto S_a by some continuous $g\colon S_b \to S_a$ *inverse* to f or in other words the maps 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 itself; and conversely:

$$\begin{tikzcd} & S_b \arrow[dr, "g"] & \\ S_a \arrow[ur, "f"] \arrow[rr, equal] & & S_a \end{tikzcd} \qquad \begin{tikzcd} & S_a \arrow[dr, "f"] & \\ S_b \arrow[ur, "g"] \arrow[rr, equal] & & S_b \end{tikzcd}$$

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 did not take the mapping property as defining sameness of structure. They took each distinct kind of structure to have its own special notion of "sameness," expressed in its own terms, and merely confirmed by existence of suitable maps. As André Weil put it, "there was great confusion: the very meaning of the word 'isomorphism' varied from one theory to another" (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 gives its own notion of isomorphism or "sameness of structure," so it implicitly defines its own corresponding notion 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."[1]



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.

[1]Geometrically this means proportionate thickness plus a measure of twisting.

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. Topology was largely created as a subject when Riemann proved deep theorems relating this complex analysis to the topology of surfaces.[2]

1.1. **Morphism-based mathematics.** 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. They talk about its *topological* structure, revealed by its place in the category **TopSurf**, and its *holomorphic* or *analytic* structure revealed by its place in **HolSurf**, and other aspects of structure revealed by its place in other categories. Relations among these aspects of structure are revealed by functors between the 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:

$$\begin{tikzcd} & A \arrow[dr, "1_A"] & \\ C \arrow[ur, "h"] \arrow[rr, "h"'] & & A \end{tikzcd} \qquad \begin{tikzcd} & A \arrow[dr, "k"] & \\ A \arrow[ur, "1_A"] \arrow[rr, "k"'] & & B \end{tikzcd}$$

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

Definition 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 $f: A \to B$.

$$\begin{tikzcd} & B \arrow[dr, "g"] & \\ A \arrow[ur, "f"] \arrow[rr, "1_A"'] & & A \end{tikzcd} \qquad \begin{tikzcd} & A \arrow[dr, "f"] & \\ B \arrow[ur, "g"] \arrow[rr, "1_B"'] & & B \end{tikzcd}$$

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.

[2] On Riemann in the origins of topology see (Ferreirós, 1999, pp. 53ff.) and Gray (1998).

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

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

Existence and uniqueness of composites:

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

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

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

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

In the 1920s and 1930s Emmy Noether shaped modern mathematics with her *homomorphism* and *isomorphism* theorems (McLarty, 2006). 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).

2. Axiomatic foundations

The Eilenberg-Mac Lane category axioms give the bare language of the insight, formalized in two-sorted first order logic in Figure 1. Less formally, the objects and arrows of any one category $\mathbf{C}$ satisfy these conditions: Every arrow f goes from a unique object A to a unique object B. Let $f\colon A \to B$ 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 $f\colon A \to B$ is the domain of $g\colon B \to C$ then they have a *composite* $gf\colon A \to C$.[3] Each object A has an *identity* arrow $1_A\colon A \to A$ defined by $f1_A = f$ and $1_B f = f$ for every $f\colon A \to B$. The last axiom says composition is associative: $(hg)f = h(gf)$.

$$\begin{tikzcd} & B \arrow[dr, "g"] & \\ A \arrow[ur, "f"] \arrow[rr, "gf"'] & & C \end{tikzcd} \qquad \begin{tikzcd} & A \arrow[dr, "f"] & \\ A \arrow[ur, "1_A"] \arrow[rr, "f1_A = f"'] & & B \end{tikzcd} \qquad \begin{tikzcd} & B \arrow[dr, "1_B"] & \\ A \arrow[ur, "f"] \arrow[rr, "1_B f = f"'] & & B \end{tikzcd} \qquad \begin{tikzcd} B \arrow[r, "hg"] \arrow[dr, "g"] & D \\ A \arrow[u, "f"] \arrow[r, "gf"'] & C \arrow[u, "h"'] \end{tikzcd}$$

[3]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.

2.1. **The Elementary Theory of the Category of Sets.** 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.[4] 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 2 gives the ETCS axioms beyond the category axioms.[5]

Intuitively, a set 1 is a singleton if and only if every set T has exactly one function to it. And a set A has as 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.

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:

$$\begin{tikzcd} & A \arrow[dr, "f"] & \\ 1 \arrow[ur, "x"] \arrow[rr, "f(x)"'] & & B \end{tikzcd} \qquad \begin{tikzcd} & A \arrow[dr, "1_A"] & \\ 1 \arrow[ur, "x"] \arrow[rr, "1_A(x)=x"'] & & B \end{tikzcd}$$

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

$$\begin{tikzcd} & A \arrow[rr, "gf"] \arrow[dr, "f"] & & C \\ 1 \arrow[ur, "x"] \arrow[rr, "f(x)"'] & & B \arrow[ur, "g"'] & \end{tikzcd}$$

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 $x \in A$:

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

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

Theorem 1. *All singletons* 1 *and* $1'$ *are isomorphic.*

Proof. If 1 and $1'$ are singletons there are functions $u : 1 \to 1'$ and $v : 1' \to 1$. 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 $s : \mathbb{N} \to \mathbb{N}$ such that: For any set T and element $x \in T$ and

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

[5] 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:

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

$$\begin{tikzcd} & T \arrow[dl, "f"'] \arrow[d, dashed, "\langle f,g\rangle"] \arrow[dr, "g"] & \\ A & A \times B \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & B \end{tikzcd}$$

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

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

$$\begin{tikzcd} T \arrow[d, "u"'] \arrow[dr, "h"] & & \\ E \arrow[r, "e"'] & A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B \end{tikzcd}$$

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

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

$$\begin{tikzcd} C \arrow[d, "\widehat{g}"'] \\ B^A \end{tikzcd} \qquad \begin{tikzcd} C \times A \arrow[r, "g"] \arrow[d, "\widehat{g}\times 1_A"'] & B \\ B^A \times A \arrow[ur, "e"'] & \end{tikzcd}$$

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

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

$$\begin{tikzcd} S \arrow[r, tail] & A \arrow[r, shift left, "\chi_i"] \arrow[r, shift right, "true_A"'] & 2 \end{tikzcd}$$

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

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

$$\begin{tikzcd} 1 \arrow[r, "0"] \arrow[dr, "x"'] & \mathbb{N} \arrow[r, "s"] \arrow[d, "u"] & \mathbb{N} \arrow[d, "u"] \\ & T \arrow[r, "f"'] & T \end{tikzcd}$$

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 2. the ETCS axioms

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 2. 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 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* $su = u's$*.*

$$\begin{tikzcd} & \mathbb{N} \arrow[r, "s"] \arrow[d, "u"] & \mathbb{N} \arrow[d, "u"] \\ 1 \arrow[ur, "0"] \arrow[dr, "0'"'] & & \\ & \mathbb{N}' \arrow[r, "s'"'] & \mathbb{N}' \end{tikzcd}$$

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:

$$\begin{tikzcd} & T \arrow[r, "f"] \arrow[d, "m"] & T \arrow[d, "m"] \\ 1 \arrow[ur, "x"] \arrow[dr, "x'"'] & & \\ & T' \arrow[r, "f'"'] & T' \end{tikzcd}$$

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

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 1 and 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 $\pi_1\colon A\times B\to A$ and $\pi_2\colon A\times B\to B$ such that for every set T with arrows to each set $f\colon T\to A$ and $g\colon T\to B$ there is a unique function $\langle f, g\rangle\colon T\to A\times B$ 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 2. 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 3. *If a set* P *and functions* $\pi'_A\colon P\to A$ *and* $\pi'_B\colon P\to B$ *also have the product property then* P *is isomorphic to* $A\times B$*; and there is exactly one isomorphism* $u\colon P\to A\times B$ *with* $\pi'_1 = \pi_1 u$ *and* $\pi'_2 = \pi_2 u$*.*

$$\begin{tikzcd} & P \arrow[dl, "\pi'_1"'] \arrow[dr, "\pi'_2"] \arrow[d, "u"] & \\ A & A\times B \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & B \end{tikzcd}$$

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

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

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

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

$$\begin{tikzcd} A \arrow[d, "f"'] & A\times B \arrow[l] \arrow[r] \arrow[d, "f\times g"'] & B \arrow[d, "g"] \\ C & C\times D \arrow[l] \arrow[r] & D \end{tikzcd}$$

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

Any one-to-one function $i\colon S\to A$ gives a *subset* of A. We often write $i\colon S\rightarrowtail A$ 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 $j\colon T\rightarrowtail A$ is included in i, written $j\subseteq i$ if there is some $h\colon S\to T$ such that $j = ih$. To put it in graphics:

$$\begin{tikzcd} 1 \arrow[rr, "y"] \arrow[dr, "x"'] & & S \arrow[dl, tail, "i"] \\ & A & \end{tikzcd} \qquad \begin{tikzcd} T \arrow[rr, "h"] \arrow[dr, tail, "j"'] & & S \arrow[dl, tail, "i"] \\ & A & \end{tikzcd}$$

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

$$\begin{tikzcd} 1 \arrow[r, "y"] \arrow[dr, "x"'] & T \arrow[r, "h"] \arrow[d, tail, "j"] & S \arrow[dl, tail, "i"] \\ & A & \end{tikzcd}$$

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 2. 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. *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 $e' = eu$.*

$$\begin{tikzcd} E' \arrow[dr, "e'"] \arrow[dd, "u"'] & & \\ & A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B \\ E \arrow[ur, "e"'] & & \end{tikzcd}$$

Proof. Obvious following the earlier proofs. $\square$

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.

$$\begin{tikzcd} S \arrow[r, tail] & A \arrow[r, shift left, "\chi_i"] \arrow[r, shift right, "true_A"'] & 2 \end{tikzcd}$$

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 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 $\chi_j i = true_A i$

$$\begin{tikzcd} S \arrow[d, "u"'] \arrow[dr, tail, "i"] & & \\ T \arrow[r, tail, "j"'] & A \arrow[r, shift left, "\chi_j"] \arrow[r, shift right, "true_A"'] & 2 \end{tikzcd}$$

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.

$$\begin{tikzcd} 1 \arrow[r, "y"] \arrow[rr, bend left, "z"] \arrow[dr, "x"'] & S \arrow[d, tail, "i"] & T \arrow[dl, tail, "j"] \\ & A & \end{tikzcd}$$ □

The last set construction axiom posits, for any sets A, B, a *function set* or *exponential* B^A with an *evaluation function* $ev\colon B^A \times A \to B$ with this property: for any set C and function $g\colon C \times A \to B$ there is a unique function $\widehat{g}\colon C \to B^A$ with $ev(\widehat{g} \times 1_A) = g$. See the diagram in Figure 2. As a basic example elements $\widehat{f} \in B^A$ correspond to functions $f\colon 1 \times A \to B$. 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 $f\colon A \to B$ 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 $f\colon A \to B$ has at least one right inverse, that is a function $g\colon B \to A$ with $gf = 1_B$.

2.1.1. *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.

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 $f\colon A \to I$ 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.

2.1.2. *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.[6] 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:[7]

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

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 section 2.1.1) 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

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**.

2.2. **The category of categories.** A *functor* $\mathbf{F}\colon\mathbf{C}\to\mathbf{D}$ maps the network of arrows of category $\mathbf{C}$ into the network of category $\mathbf{D}$. Each object A of $\mathbf{C}$ is assigned an object $\mathbf{F}A$ of $\mathbf{D}$. Each arrow $f\colon A\to A'$ of $\mathbf{C}$ is assigned an arrow $\mathbf{F}f\colon\mathbf{F}A\to\mathbf{F}A'$ of $\mathbf{D}$ preserving the domain and codomain as shown. A functor preserves identity arrows in that for each object A of $\mathbf{C}$

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

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

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

And it maps each composition triangle in $\mathbf{C}$ to one in $\mathbf{D}$:

$$\begin{tikzcd} & A' \arrow[dr, "g"] & \\ A \arrow[ur, "f"] \arrow[rr, "gf"'] & & A'' \end{tikzcd} \qquad \begin{tikzcd} & \mathbf{F}A' \arrow[dr, "\mathbf{F}g"] & \\ \mathbf{F}A \arrow[ur, "\mathbf{F}f"] \arrow[rr, "\mathbf{F}(gf)=(\mathbf{F}g)(\mathbf{F}f)"'] & & \mathbf{F}A'' \end{tikzcd}$$

Depending on the choice of foundation it is either an axiom or a obvious theorem that each category $\mathbf{C}$ has an identity functor $1_{\mathbf{C}}\colon \mathbf{C}\to\mathbf{C}$, while functors $\mathbf{F}\colon\mathbf{C}\to\mathbf{D}$ and $\mathbf{G}\colon\mathbf{D}\to\mathbf{E}$ compose to give a functor $\mathbf{GF}\colon\mathbf{C}\to\mathbf{E}$. Composition is associative, so categories and functors themselves form a category—although depending on foundations this category may not exist! Let us be naive for the moment.

Functors are everywhere.[8] Riemann implicitly had a functor

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

inserting the holomorphic category of Riemann surfaces into the topological category. This functor takes each holomorphic map $f\colon S_a\to S_b$ of surfaces to itself as a continuous map. A further forgetful functor $\mathbf{TopSurf}\to\mathbf{Set}$ takes each surface to its set of points and each continuous map $f\colon S_a\to S_b$ to itself as a mere function between sets. 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

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

from the category of topological spaces and continuous maps to the category of Abelian groups. Each continuous map of spaces $f\colon T\to T'$ induces a group homomorphism $H_n(f)\colon H_n(T)\to H_n(T')$ so that the most relevant topological information is concentrated into a simpler algebraic form.[9] Functors are too common and too varied to need or allow a comprehensive survey here. Naively, then, categories and functors form a category, or if you prefer a *meta-category.* Here we will not look for one absolute category of all categories (any more than set theorists try to find one absolute universe of all sets) but will first intuitively and then axiomatically describe a category of categories as a foundation for mathematics.

2.2.1. *Categorical category theory.* Lawvere (1963) first systematically applied the central insight itself to functors.[10] Categories, including a category $\mathbf{Set}$ 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

---

[8]See examples in Mac Lane (1986, Ch. XI.9) and the more advanced (Mac Lane, 1998) which advises more specifically "adjoint functors arise everywhere" [p. vii].

[9]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).

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

naively look like this:

$$
\begin{array}{ccc}
0 & \begin{tikzcd} 0 \arrow[r, "\alpha"] & 1 \end{tikzcd} & \begin{tikzcd} & 1 \arrow[dr, "\beta"] & \\ 0 \arrow[ur, "\alpha"] \arrow[rr, "\gamma"'] & & 2 \end{tikzcd} \\
\mathbf{1} & \mathbf{2} & \mathbf{3}
\end{array}
$$

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:

$$
\begin{tikzcd} & \mathbf{2} \arrow[dr, "f"] & \\ \mathbf{1} \arrow[ur, "0"] \arrow[rr, "f_0"'] & & \mathbf{C} \end{tikzcd}
\qquad
\begin{tikzcd} & \mathbf{2} \arrow[dr, "f"] & \\ \mathbf{1} \arrow[ur, "1"] \arrow[rr, "f_1"'] & & \mathbf{C} \end{tikzcd}
$$

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

$$
\begin{tikzcd} & \mathbf{1} \arrow[dr, "A"] & \\ \mathbf{2} \arrow[ur] \arrow[rr, "1_A"'] & & \mathbf{C} \end{tikzcd}
$$

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:

$$
\begin{tikzcd} & \mathbf{C} \arrow[dr, "\mathbf{F}"] & \\ \mathbf{1} \arrow[ur, "A"] \arrow[rr, "\mathbf{F}A"'] & & \mathbf{D} \end{tikzcd}
\qquad
\begin{tikzcd} & \mathbf{C} \arrow[dr, "\mathbf{F}"] & \\ \mathbf{2} \arrow[ur, "f"] \arrow[rr, "\mathbf{F}f"'] & & \mathbf{D} \end{tikzcd}
$$

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

$$
\begin{tikzcd} & 1 \arrow[r, "\mathbf{F}A"] \arrow[dr, "A"] & D \\ 2 \arrow[ur] \arrow[r, "1_A"'] & C \arrow[ur, "\mathbf{F}"'] & \end{tikzcd}
$$

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 $f:\mathbf{2}\to\mathbf{C}$ and $g:\mathbf{2}\to\mathbf{C}$ compose as arrows of $\mathbf{C}$ if and only if $f_1 = g_0$:

$$\begin{tikzcd}
& \mathbf{2} \arrow[dr, "f"] & \\
\mathbf{1} \arrow[ur, "1"] \arrow[dr, "0"'] & & \mathbf{C} \\
& \mathbf{2} \arrow[ur, "g"'] &
\end{tikzcd}$$

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 $t:\mathbf{3}\to\mathbf{C}$ with $t\alpha = f$ and $t\beta = g$:

$$\begin{tikzcd}
\mathbf{2} \arrow[dr, "\alpha"'] \arrow[drr, bend left, "f"] & & \\
& \mathbf{3} \arrow[r, "t"] & \mathbf{C} \\
\mathbf{2} \arrow[ur, "\beta"] \arrow[urr, bend right, "g"'] & &
\end{tikzcd}$$

Then the composite of f and g in $\mathbf{C}$ can be defined as the third side $t\gamma:\mathbf{2}\to\mathbf{C}$. 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 $t\beta = g$:

$$\begin{tikzcd}
& \mathbf{2} \arrow[dr, "\alpha"'] \arrow[drr, bend left, "f"] & & \\
\mathbf{1} \arrow[ur, "1"] \arrow[dr, "0"'] & & \mathbf{3} \arrow[r, "t"] & \mathbf{C} \\
& \mathbf{2} \arrow[ur, "\beta"] \arrow[urr, bend right, "g"'] & &
\end{tikzcd}$$

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.[11]

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 $\mathbf{Set}$ 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:

$$\begin{tikzcd}
& \mathbf{1} \arrow[dr, "0"] & \\
\mathbf{2} \arrow[ur] \arrow[rr, "1_0"'] & & \mathbf{2}
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{2} \arrow[r, "1_{\mathbf{2}}"] & \mathbf{2}
\end{tikzcd}
\qquad
\begin{tikzcd}
& \mathbf{1} \arrow[dr, "1"] & \\
\mathbf{2} \arrow[ur] \arrow[rr, "1_1"'] & & \mathbf{2}
\end{tikzcd}$$

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 $\mathbf{2}$.[12]

[11] 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 $\mathbf{C}$.

[12] 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).

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 $f\colon \mathbf{2} \to \mathbf{C}$ 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 $x\colon 1 \to S$. Categories have extensionality with regard to the arrow category $\mathbf{2}$: distinct functors between categories $\mathbf{C} \to \mathbf{D}$ differ on some arrow $f\colon \mathbf{2} \to \mathbf{C}$.

A category is *discrete* if every arrow $f\colon A \to B$ in it is an identity: that is $A = B$ and $f = 1_A$. Notice that a functor $\mathbf{F}\colon \mathbf{D} \to \mathbf{D}'$ 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 section: 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**.

2.2.2. *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. For the rest of this section 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 1 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 1, 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 3.

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:

$$\begin{tikzcd} & \mathbf{1} \arrow[dr, "0"] & \\ \mathbf{2} \arrow[ur] \arrow[rr, "1_0"'] & & \mathbf{2} \end{tikzcd} \qquad \begin{tikzcd} \mathbf{2} \arrow[r, "1_{\mathbf{2}}"] & \mathbf{2} \end{tikzcd} \qquad \begin{tikzcd} & \mathbf{1} \arrow[dr, "1"] & \\ \mathbf{2} \arrow[ur] \arrow[rr, "1_1"'] & & \mathbf{2} \end{tikzcd}$$

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

$$\begin{tikzcd} & \mathbf{2} \arrow[dr, "\alpha"] & \\ \mathbf{1} \arrow[ur, "1"] \arrow[dr, "0"'] & & \mathbf{3} \\ & \mathbf{2} \arrow[ur, "\beta"'] & \end{tikzcd}$$

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

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

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

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

By the axiom on $\mathbf{3}$, any arrows $f\colon\mathbf{2}\to\mathbf{A}$ and $g\colon\mathbf{2}\to\mathbf{A}$ with $f_1 = g_0$ determine a unique $t\colon\mathbf{3}\to\mathbf{A}$ 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.[13]

Figure 4 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:

$$\begin{tikzcd} & \mathbf{T} \arrow[dl, "\mathbf{F}"'] \arrow[d, dashed, "\langle\mathbf{F},\mathbf{G}\rangle"] \arrow[dr, "\mathbf{G}"] & \\ \mathbf{A} & \mathbf{A}\times\mathbf{B} \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & \mathbf{B} \end{tikzcd} \qquad \begin{tikzcd} & \mathbf{T} & \\ \mathbf{A} \arrow[ur, "\mathbf{F}"] \arrow[r, "i_1"'] & \mathbf{A}+\mathbf{B} \arrow[u, "\binom{\mathbf{F}}{\mathbf{G}}"] & \mathbf{B} \arrow[ul, "\mathbf{G}"'] \arrow[l, "i_2"] \end{tikzcd}$$

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

$$\begin{tikzcd} \mathbf{T} \arrow[d, "\mathbf{u}"'] \arrow[dr, "\mathbf{H}"] & & \\ \mathbf{E} \arrow[r, "\mathbf{e}"'] & \mathbf{A} \arrow[r, shift left, "\mathbf{F}"] \arrow[r, shift right, "\mathbf{G}"'] & \mathbf{B} \end{tikzcd} \qquad \begin{tikzcd} & & \mathbf{T} \\ \mathbf{B} \arrow[r, shift left, "\mathbf{F}"] \arrow[r, shift right, "\mathbf{G}"'] & \mathbf{A} \arrow[ur, "\mathbf{H}"] \arrow[r, "\mathbf{q}"'] & \mathbf{Q} \arrow[u, "\mathbf{u}"'] \end{tikzcd}$$

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

$$\begin{tikzcd} \mathbf{C} \arrow[d, "\widehat{\mathbf{G}}"'] \\ \mathbf{B}^{\mathbf{A}} \end{tikzcd} \qquad \begin{tikzcd} \mathbf{C}\times\mathbf{A} \arrow[r, "\mathbf{G}"] \arrow[d, "\widehat{\mathbf{G}}\times 1_{\mathbf{A}}"'] & \mathbf{B} \\ \mathbf{B}^{\mathbf{A}}\times\mathbf{A} \arrow[ur, "\mathbf{e}"'] & \end{tikzcd}$$

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

$$\begin{tikzcd} \mathbf{1} \arrow[r, "\mathbf{0}"] \arrow[dr, "\mathbf{x}"'] & \mathbf{N} \arrow[r, "\mathbf{S}"] \arrow[d, "\mathbf{u}"] & \mathbf{N} \arrow[d, "\mathbf{u}"] \\ & \mathbf{A} \arrow[r, "\mathbf{F}"'] & \mathbf{A} \end{tikzcd}$$

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 $\mathbf{F}\mathbf{H} = 1_{\mathbf{D}}$.

FIGURE 4. The CCAF construction axioms

[13] McLarty (1991a). The **Set** axiom in Figure 5 replaces several axioms in that paper.

The axioms in Figure 5 posit a category **Set** of sets with a fullness condition. The condition uses categories constructed from **Set** by axioms in Figure 4.[14] For each set $I : \mathbf{1} \to \mathbf{Set}$ 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 **Set**. All its arrows are identity arrows.

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 5. 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 $\mathbf{A}$, does. In ETCS a subcategory of $\mathbf{A}$ 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 $\mathbf{A}$ to those of a category $\mathbf{B}$ which intuitively ought to define a functor $\mathbf{A} \to \mathbf{B}$, does.[15] And so coequalizers of categories have the properties they intuitively ought to.[16] 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 $\mathbf{A}$ (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 $\mathbf{N}, \mathbf{0}, \mathbf{S}$.

2.2.3. *A sample construction.* The category **Iter** was described intuitively in section 2.1: 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 $1 : \mathbf{1} \to \mathbf{Set}$, 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

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

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

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

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

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

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

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

By symmetry of i and i' in δ, it follows that k is an isomorphism of $\mathbf{B}$ and $\mathbf{B}'$.

FIGURE 6. The CCAF separation axiom scheme

it corresponds to an object of the product category $\mathbf{Set^2}\times\mathbf{Set^2}$. 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

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

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 $\mathbf{Set^2}\times\mathbf{Set^2}$. 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 $\mathbf{Set^2}\times\mathbf{Set^2}$ *is* a pair of functions. Rather we use the constructed inclusion functor, plus the two projections

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

to *determine* a pair of functions for each object of **Iter**. And of course $\mathbf{Set^2}$ 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.

3. ONTOLOGY

3.1. **Structures.** On this foundation a *structure* is any object in a category. This includes the usual structured sets as for example section 2.2.3 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 3–6 posit the existence of certain categories and functors, including one category **Set** that satisfies the ETCS axioms in Figure 2, 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.[17]

(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.[18]

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 section 4 below). Tarski's heritage, celebrated at these congresses, urges us to make such things precise.

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.[19] 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 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 3 on indistinguishability rules out constants referring to such

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

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

[19] 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.

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.

3.2. **Theorem schemes.** Write

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

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

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

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 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 $\mathcal{Q}$:*

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

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 $A\colon \mathbf{1} \rightarrow \mathbf{C}$. 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 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 $\mathcal{P}$:*

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

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 $i\colon S \rightarrow S$ 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 $x\colon 1 \rightarrow S$ in **Set** and so it is a particular kind of functor $f\colon \mathbf{2} \rightarrow \mathbf{Set}$:

Theorem scheme 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 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.

4. Naturalism

Philosophy of mathematics ought to reflect the values of mathematical practice itself, along something like 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 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 but we are 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).

The unification impacts foundations on the basic question of what exists. Maddy says 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 (Maddy, 2007, p. 363)

But innumerable perfectly orthodox textbooks and research papers affirm that the category of sets **Set** exists, and the category **Grp** of groups, and much more. These cannot exist in either Zermelo Fraenkel set theory 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 for the existence of various large categories. Lawvere (1963, 1966) did that with CCAF.

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*.[20] 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)

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.

[20] 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, 2007) describes briefly for philosophers.

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 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:

$$\begin{tikzcd} & T \arrow[dl, "f"'] \arrow[d, dashed, "\langle f,g\rangle"] \arrow[dr, "g"] & \\ M & M\times N \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & N \end{tikzcd}$$

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."[21] Yet ETCS sets are rigorously composed of mere units and they remain so in CCAF.

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

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)[22]

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 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.

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:

[22]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?

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.

5. Conclusion

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 (2007). The central insight into structure is not the whole of mathematics. It pervades mathematics from research to textbooks to foundations, though, and it is this: What matters about structures is their maps.

References

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.

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. (2007). Categories for the working mathematician: Making the impossible possible. *Synthese*. In press, currently in SpringerLink Online First.

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

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

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.

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

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.

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

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

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

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

Lawvere, F. W. (1963). *Functorial Semantics of Algebraic Theories*. PhD thesis, Columbia University. Republished with commentary by the author 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. Republished with commentary by the author and Colin McLarty in: *Reprints in Theory and Applications of Categories*, No. 11 (2005) pp. 1–35, available 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.

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

Mac Lane, S. (1998). *Categories for the Working Mathematician*. Springer-Verlag, New York, 2nd edition.

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

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

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

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.

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. (2006). 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.

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

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

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. 1–45.

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

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