

Structuralism, Invariance, and Univalence*

Steve Awodey

March 4, 2014

Abstract

The recent discovery of an interpretation of constructive type theory into abstract homotopy theory suggests a new approach to the foundations of mathematics with intrinsic geometric content and a computational implementation. Voevodsky has proposed such a program, including a new axiom with both geometric and logical significance: the *Univalence Axiom*. It captures the familiar aspect of informal mathematical practice according to which one can identify isomorphic objects. While it is incompatible with conventional foundations, it is a powerful addition to homotopy type theory. It also gives the new system of foundations a distinctly structural character.

Recent advances in foundations of mathematics have led to some developments that are significant for the philosophy of mathematics, particularly structuralism. Specifically, the discovery of an interpretation of Martin-Löf's constructive type theory into abstract homotopy theory [3] suggests a new approach to the foundations of mathematics, one with both intrinsic geometric content and a computational implementation [5]. Leading homotopy theorist Vladimir Voevodsky has proposed an ambitious new program of foundations on this basis, including a new axiom with both geometric and logical significance: the *Univalence Axiom* [4]. It captures the familiar aspect of informal mathematical practice according to which one can identify isomorphic objects. While it is incompatible with conventional foundations, it is a powerful addition to the framework of homotopical type theory.

*Thanks to Peter Aczel for many discussions on the subject of this paper, and to the Munich Center for Mathematical Philosophy, where this work was done and presented.

1 The Principle of Structuralism

The following statement may be called the *Principle of Structuralism*:

Isomorphic objects are identical. (PS)

From one perspective, this captures a principle of reasoning embodied in everyday mathematical practice:

- the fundamental group of the circle $\pi_1(S^1)$ is isomorphic to the group \mathbb{Z} of integers, so for the purposes of group theory, these are *the same* group, namely the free group on one generator, $F(1)$.
- The Cauchy reals are isomorphic to the Dedekind reals, so as far as analysis is concerned, these are *the same* number field, \mathbb{R} .
- the unit interval $[0, 1]$ is homeomorphic to the closed interval $[0, 2]$, so topologically, these are *the same* space, say I .

Within a mathematical theory, theorem, or proof, it makes no practical difference which of two “isomorphic copies” are used, and so they can be treated as *the same* mathematical object for all practical purposes. This common practice is even sometimes referred to light-heartedly as “abuse of notation,” and mathematicians have developed a sort of systematic sloppiness to help them implement this principle, which is quite useful in practice, despite being literally false. It is, namely, incompatible with conventional foundations of mathematics in set theory.

Indeed, from a more literal point view, it is just false that isomorphic objects are identical. The fundamental group consists of (equivalence classes of) paths, not numbers. The Cauchy reals are sequences of rationals, i.e. functions $\mathbb{N} \rightarrow \mathbb{Q}$, and the Dedekind reals are subsets of rationals, etc.

Mathematical objects are often constructed out of other ones, and thus also have some residual structure resulting from that construction, in addition to whatever structure they may have as objects of interest, i.e. the real number field actually consisting of functions or of subsets. So there is a clear sense in which (PS) is simply false.

What if we try to be generous and read “identity” in a weaker way:

“ A is identical to B ” means “ A and B have all the same properties”.

Here we may want to restrict the properties that can occur on the right, in order for this to really be a weaker condition. So we now have the reformulation:

$$A \cong B \Rightarrow \text{for all relevant properties } P, P(A) \Rightarrow P(B) \quad (\text{PS}')$$

The relevant properties will be the ones pertaining to the subject matter: group-theoretic, topological, etc. However, consider for example perfectly reasonable set-theoretic properties like:

$$\cap X \in X, \quad \text{or} \quad \emptyset \in X.$$

Such mundane properties of sets (and so of any structures built from sets) do not “respect isomorphism” in the sense stated in (PS’). So the task of precisely determining the “relevant” properties is evidently no simple matter.

Our leading question may now be formulated as follows:

Is there a sense in which (PS) can be held true?

Such a reading would legitimize much mathematical practice and support a structuralist point of view. Let us consider the terms of (PS) in turn.

2 Isomorphism

What does it mean to say that two things are isomorphic? Some might be tempted to say:

$$A \cong B \Leftrightarrow A \text{ and } B \text{ have the same structure.} \quad (1)$$

And that would be true.

But it is not the *definition* of isomorphism, because it presumes the notion of structure and an identity criterion for it, and that is putting the cart before the horse. *Structure* is like *color* or *the direction of a line* or *number*: it is an abstract concept. And as Frege wisely taught us, it is determined by an abstraction principle, in this case:

$$\text{str}(A) = \text{str}(B) \Leftrightarrow A \cong B. \quad (\text{DS})$$

The structure of A is the same as the structure of B just in case A and B are isomorphic. So (1) is indeed true, but it’s the definition of “structure”

in terms of isomorphism, not the other way around. (Of course, such an informal “definition” of the notion of structure would require a more explicit formal setting to actually be of use; we shall take a different course to arrive at the same result.)

So what is the definition of “isomorphism”? Two things A and B are *isomorphic*, written $A \cong B$, if there are structure-preserving maps

$$f : A \rightarrow B \quad \text{and} \quad g : B \rightarrow A$$

such that

$$g \circ f = 1_A \quad \text{and} \quad f \circ g = 1_B.$$

This standard definition of isomorphism makes reference to “structure-preserving maps”, “composition” \circ , and “identity” maps 1_A , but these are primitive concepts in category theory, and so need not be further defined. Note that “isomorphism” is always relative to a given category, which determines a kind of structure via a primitive notion of “structure-preserving maps”, and not the other way around.

Sometimes structures are built up from other ones, like groups, graphs, spaces. These are sets equipped with operations, relations, etc. The notion of a structure-preserving map, or *morphism*, is then defined in terms of maps in the underlying category of sets (i.e. functions) that preserve the operations, relations, etc. — i.e. “homomorphisms,” in the usual sense. But there are lots of other examples where the notion of morphism is given directly, and that then determines the corresponding notion of “structure” — e.g. “differentiable structure” or “smooth structure”.¹

¹This view of the concept of mathematical structure as determined via category theory — a category determines a species of structure, rather than the other way around — was already presented in [1]. It results in a notion of structure that differs radically from that presumed in some of the philosophical literature, e.g. [9, 10]. Here a structure is determined *externally*, as it were, by its mappings to and from other objects of the same kind, rather than *internally*, in terms of relations and operations on elements. This solves Benaceraff-style problems directly by rejecting the idea that “mathematical objects” are the elements of structured sets (e.g. particular numbers) in favor of the view that they are structures (e.g. the system of all numbers, together with 0 and the operation of successor).

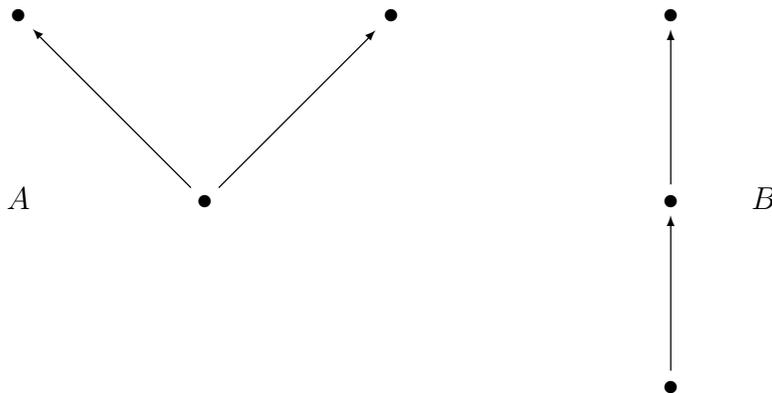
3 Objects

If two objects are isomorphic, they stay that way if we forget some of the structure. For example, isomorphic groups are also isomorphic as sets,

$$G \cong H \Rightarrow |G| \cong |H|.$$

Therefore, one can *distinguish* different structures by taking them to non-isomorphic underlying sets.

More generally, any functor (not just a forgetful one) preserves isomorphisms, so we can distinguish non-isomorphic structures by taking them to non-isomorphic objects by some functor. For example, consider the two simple partial orders:



Are they isomorphic? Without even thinking about what order-preserving maps back and forth there may be, we can simply count the number of such maps from the simple partial order $I = (\cdot \rightarrow \cdot)$, and we see immediately that A has 5, and B has 6. Since taking the set $\mathbf{Hom}(I, X)$ of maps from any fixed object, here I , is always a functor of X , it follows immediately that A and B cannot be isomorphic.

A property $P(X)$ like this is called an *invariant*:

$$A \cong B \ \& \ P(A) \Rightarrow P(B)$$

It is a property that respects isomorphism, i.e. a “structural property”. If in our (restricted) Principle of Structuralism (PS’) we were to restrict attention to only structural properties P in this sense, we would get the statement:

$$A \cong B \Rightarrow \text{for all structural properties } P, P(A) \Rightarrow P(B).$$

That is, “isomorphic objects have all the same structural properties”. Now, this statement is trivially true, because of the way we’ve set up the definitions, but it does indeed capture the way in which (PS) is actually used in practice. Namely:

If $A \cong B$ and $P(X)$ is any structural property such that $P(A)$, then also $P(B)$.

But in order for this to actually be *useful* in practice, we obviously need to be able to recognize that $P(X)$ is structural, without going through the trouble of proving in general that it respects isomorphisms! Let us consider one common way of doing just that.

Suppose that we have one or more functors from the category \mathcal{S} of structures of a given kind to the category of sets,

$$F : \mathcal{S} \rightarrow \mathbf{Sets}.$$

Then the identity of structures – i.e. isomorphism in \mathcal{S} – can be tested via the functors F in terms of the identity of the structure of the underlying “sets”, i.e. set isomorphisms. This was the situation in our example above of counting the number of homomorphisms from a fixed object: there our F was the functor $\mathbf{Hom}(I, -)$. A typical kind of problem asks whether there are enough such “invariants” F of a certain kind to determine any given structure in \mathcal{S} : e.g. is the homotopy type of a space determined by its homotopy groups? is a formula provable if it’s true in all models? etc.

Now that we have reduced identity of structures in general to the question of isomorphism of some “base category”, like \mathbf{Sets} , we just need to know: what are the invariant properties of \mathbf{Sets} ? We already know that many sentences of conventional set theory express properties of sets that are *not* invariant, like $\emptyset \in X$. One way to obtain invariant properties is by restricting the methods used to specify them to only combinations of ones that are already known to be structural. As we have seen, category-theoretic methods are always structural; but another large and useful class of structural properties are those that can be defined in a foundational system of *type theory*, rather than set theory.²

²These two different kinds of invariants are actually closely related, as explained by categorical logic.

The basic operations of what is called *constructive type theory* [7], starting with a basic object or “type of individuals” X , are as follows:

$$0, 1, A + B, A \times B, A \rightarrow B, \Sigma_{x:A} B(x), \Pi_{x:A} B(x), \text{Id}_A(x, y).$$

and these correspond to the logical propositions:

$$\perp, \top, A \vee B, A \wedge B, A \Rightarrow B, \exists x : A. B(x), \forall x : A. B(x), x =_A y.$$

The correspondence – called “Curry-Howard” or “Propositions-as-Types” – is given by:

$$\text{proof} : \text{Proposition} \approx \text{term} : \text{Type}$$

That is, the proofs of a proposition are the terms of the corresponding type. We shall return to this idea below.

The system of type theory has the important property that *any* definable property of objects is invariant. Thus if $P(X)$ is any type that is definable in the system over a basic type X , then the following inference is derivable:

$$\frac{A \cong B \quad P(A)}{P(B)}$$

The proof is a straightforward induction on the construction of $P(X)$. Let us record this as the following *Principle of Invariance* for type theory:³

$$\text{All definable properties are isomorphism invariant.} \quad (\text{PI})$$

4 Identity

What is meant by “the objects A and B are identical”? In set theory it means $\forall x. x \in A \Leftrightarrow x \in B$, but we have already opted out of set theory, for good reasons. In *impredicative* type theory, one can define identity by Leibniz’s Law as:

$$A = B := \forall P. P(A) \Rightarrow P(B),$$

But this is identity between terms $A, B : X$ of some common type X , and the universal quantifier $\forall P$ is over the type $\mathcal{P}(X)$ of all properties on (or, if you prefer, subsets of) X . So, more explicitly:

$$A =_X B := \forall P : \mathcal{P}(X). P(A) \Rightarrow P(B),$$

³The importance of this principle has also been emphasized by Makkai [8].

which is not the relation we are looking for, i.e. the one between *types* X and Y .

In constructive type theory there is a primitive identity relation $\text{ld}_X(x, y)$ on the terms of each type X , and its rules do allow the expected inference,

$$\frac{\text{ld}_X(a, b) \quad P(a)}{P(b)}$$

for any predicate $P(x)$ on X , giving (half of) the effect of the Leibnizian definition.

In order to be able to reason about identity of *types*, one can add a universe U of all (small) types; this then has its identity relation satisfying:

$$\frac{\text{ld}_U(A, B) \quad P(A)}{P(B)}$$

for any definable property $P(X)$ of types.

Now that we have found a way to reason about identity of types, we can ask how it is related to isomorphism. We can start by asking whether this extension by a universe is still compatible with the Principle of Invariance. Before we added U , we observed that all definable properties $P(X)$ are invariant, in the sense that:

$$\frac{A \cong B \quad P(A)}{P(B)}$$

If this inference also held for properties P involving U , we could set

$$P(X) := \text{ld}_U(A, X),$$

and infer from $\text{ld}_U(A, A)$ that:

$$A \cong B \Rightarrow \text{ld}_U(A, B)$$

i.e. the Principle of Structuralism.

Thus in the full system of type theory, the Principle of Invariance implies the Principle of Structuralism. To put it more simply, if in the extended system of type theory with a universe it is still the case that all definable properties are isomorphism invariant, then in particular isomorphic objects are identical.

Is that really possible?

5 What does “are” mean?

In type theory, every proposition determines a type (namely, the type of proofs of the proposition), and every type can be regarded as a proposition (namely, the proposition that the type has terms). For instance, corresponding to the proposition $A \cong B$ we have the type of isomorphisms between A and B :

$$“A \cong B” \approx \text{Iso}(A, B)$$

In order to prove that $A \cong B$, one constructs a term of type $\text{Iso}(A, B)$, which is exactly an isomorphism between A and B .

Similarly, associated to the identity type $\text{Id}_U(A, B)$ there is the proposition $A =_U B$ that A and B are identical (small) types,

$$“A =_U B” \approx \text{Id}_U(A, B).$$

The object $\text{Id}_U(A, B)$ can be regarded as the type of “proofs that A is identical B ”, or the type of “identifications of A and B ”. (There is a neat geometric interpretation of this type, that we unfortunately cannot go into here; see [5]). For the remainder of this essay, we shall follow the usual type-theoretic practice of simply *identifying propositions and types*, and use the more familiar notation $A \cong B$ and $A =_U B$ exclusively.

Now, it is easily shown that there is always a map,

$$(A =_U B) \rightarrow (A \cong B),$$

since the relation of isomorphism is reflexive. The Univalence Axiom implies that if A and B are sets, then this map is itself an isomorphism,

$$(A =_U B) \cong (A \cong B).$$

Thus in particular, there is a map coming back,

$$(A \cong B) \rightarrow (A =_U B),$$

which may reasonably be read “isomorphic objects are identical” – that is to say, the Principle of Structuralism. Indeed, this is the inference that we just doubted was even possible.

Let us immediately say that Voevodsky’s Univalence Axiom itself actually has a more general form, namely that *identity of objects is equivalent to equivalence*,

$$(A =_U B) \simeq (A \simeq B), \tag{UA}$$

where the notion of “equivalence” is a broad generalization of isomorphism that subsumes homotopy equivalence of spaces, categorical equivalence of (higher-) groupoids, isomorphism of sets and set-based structures like groups and rings, and logical equivalence of propositions.

Since the rules of identity permit substitution of identicals, one consequence of (UA) is the schema that we called the Principle of Invariance,

$$\frac{A \simeq B \quad P(A)}{P(B)}$$

but now this holds for *all* objects A and B and *all* properties $P(X)$. Thus even in this extended system with a universe and the Univalence Axiom, it still holds that *all definable properties are structural*. Indeed, preserving the Principle of Invariance *requires* that we also add the Univalence Axiom when we add a universe. In this sense, (UA) is a very natural assumption, since it determines the (otherwise underdetermined) identity relation on the universe in a way that preserves the character of the system without a universe. Rather than viewing it as identifying equivalent objects, and thus *collapsing* distinct objects, it is more useful to regard it as *expanding* the notion of identity to that of equivalence. For mathematical purposes, this is the sharpest notion of identity available; the question whether two equivalent mathematical objects are “really” identical in some stronger, non-logical sense, is thus outside of mathematics.

As fascinating as this result may be from a philosophical perspective, it should be said that it is not the only motivation for the Univalence Axiom, or even the primary one. Univalence is not a philosophical claim but rather a working axiom, with many important mathematical consequences and applications, in a new system of foundations of mathematics [4]. For example, it is used heavily in the recently-discovered “logical” calculations of some of the homotopy groups of spheres (see [5]).

In sum, the Univalence Axiom is a new axiom of logic that not only makes sense of, but actually implies, the Principle of Structuralism with which we began:

$$\text{Isomorphic objects are identical.} \quad (\text{PS})$$

This seems quite radical from a conventional foundational point of view, and it is indeed incompatible with classical foundations in set theory. It is, however, fully compatible and even natural within a type-theoretic foundation,

and that is part of this remarkable new insight.⁴

6 Background

From a foundational perspective, the Univalence Axiom is certainly radical and unexpected, but it is not entirely without precedent.

The first edition of *Principia Mathematica* used an intensional type theory, but the axiom of reducibility implied that every function has an extensionally equivalent predicative replacement. This also had the effect of spoiling the interpretation of functions as expressions – open sentences – and the substitutional interpretation of quantification, what may be called the “syntactic interpretation”, as favored by Russell on the days when he was a constructivist. In the second edition, Russell states a new principle that he attributes to Wittgenstein: “a function can only occur in a proposition through its values”, and he says that this justifies an axiom of extensionality, thus doing at least some of the work of the axiom of reducibility. What is going on here is that Wittgenstein has noticed that all functions that actually occur, i.e. that can be explicitly defined, are in fact extensional, and so one can consistently add the axiom of extensionality without destroying the syntactic interpretation.

We have here a somewhat similar move: the Univalence Axiom is in some sense a very general extensionality principle. Indeed, it even implies the usual extensionality laws for propositions and propositional functions:

$$\begin{aligned} p \leftrightarrow q &\rightarrow p = q \\ \forall x(fx \leftrightarrow gx) &\rightarrow f = g \end{aligned}$$

Since nothing in syntax violates UA, we can add it to the system and still maintain the good properties of syntax, like invariance under isomorphism.

Rudolf Carnap was one of the first people to observe that the properties definable in the theory of types are invariant under isomorphisms (he did so in pursuit of his ill-fated *Gabelbarkeitssatz*). Tarski later proposed this condition as a sort of explication of the concept of a “logical notion”. The idea was to generalize Felix Klein’s program from geometric to “arbitrary” transformations, in order to achieve the most general notion of an “invariant”, which would then be a *logical notion*.

⁴For further information on the Univalent Foundations of Mathematics, see [6, 11, 4, 2].

The consistency of the Univalence Axiom shows that the entire system of type theory is in fact invariant under the even more general notion of *homotopy equivalence* — i.e. “same shape” — which is a much larger class of maps than Tarski’s bijections of sets. The Univalence Axiom takes this basic insight,

$$A \simeq B \ \& \ P(A) \Rightarrow P(B),$$

i.e. “all logical properties are invariant”, and turns it into a logical axiom,

$$(A \simeq B) \simeq (A = B).$$

i.e. “logical identity is equivalent to equivalence”.

Finally, observe that, as an informal consequence of (UA), together with the very definition of “structure” (DS), we have that two mathematical objects are identical if and only if they have the same structure:

$$\text{str}(A) = \text{str}(B) \Leftrightarrow A = B.$$

In other words, mathematical objects simply *are* structures. Could there be a stronger formulation of structuralism?

References

- [1] S. Awodey, *Structure in mathematics and logic: A categorical perspective*, *Philosophia Mathematica* 3, vol. 4 (1996), 209–237.
- [2] S. Awodey, *Type theory and homotopy*, *Epistemology versus Ontology: Essays on the Philosophy and Foundations of Mathematics in Honour of Per Martin-Löf*, P. Dybjer et al. (ed.s), Springer 2012, 183–201.
- [3] S. Awodey and M. A. Warren, *Homotopy theoretic models of identity types*, *Mathematical Proceedings of the Cambridge Philosophical Society* 146 (2009), no. 1, 45–55.
- [4] S. Awodey, Á. Pelayo and M. A. Warren, *Voevodsky’s Univalence Axiom in homotopy type theory*, *Notices of the American Mathematical Society* (forthcoming).
- [5] *Homotopy Type Theory: Univalent Foundations of Mathematics*, The Univalent Foundations Program, Institute for Advanced Study, 2013. Available online at <http://homotopytypetheory.org/book>
- [6] *Homotopy Type Theory* (website and blog), online at <http://homotopytypetheory.org>
- [7] P. Martin-Löf, *An intuitionistic theory of types: predicative part*, *Logic Colloquium ’73* (Bristol, 1973), *Studies in Logic and the Foundations of Mathematics*, Vol. 80. North-Holland, Amsterdam, 1975, 73–118.
- [8] Makkai, M., *Towards a Categorical Foundation of Mathematics*, *Logic Colloquium ’95*, *Lecture Notes in Logic* (Volume 11), Berlin: Springer, 1998, 153–190.
- [9] Resnik, M., *Mathematics as a Science of Patterns*, Oxford: Clarendon Press, 1997.
- [10] Shapiro, S., *Philosophy of Mathematics: Structure and Ontology*, Oxford: Oxford University Press, 1997.
- [11] *Univalent Foundations of Mathematics Wiki* (UF-IAS-2012), online at <http://uf-ias-2012.wikispaces.com>