Notes on the philosophy of mathematics

John T. Baldwin
Department of Mathematics, Statistics and Computer Science
University of Illinois at Chicago

September 27, 1997

The following are some notes which may eventually develop into a paper either in philosophy or pseudo-philosophy. In the meantime comments are welcomed. Certain recepients will recognize the examples chosen specifically to tease them. For emphasis, I have deliberately taken some positions more provacatively than is actually tenable.

January 28,1987 (slightly revised August 1988)

1 A contemporary myth

The contemporary philosophy of mathematics is built upon a myth. The myth asserts that all of mathematics can be reduced to certain foundations. The philosophy of mathematics is thus reduced to the analysis of these foundations - to the unholy triumverate of logicism, formalism, and intuitionism. The substantial post-Godel developments of mathematical logic are dismissed as mere 'technical mathematics' of no philosophical significance.

We object to both the myth and its corollary dismissal. We first argue, in common with many recent writers, that although a reduction of mathematics to certain basic propositions can take place this reduction loses much of the meaning of mathematics. In particular, a number of important philosphical questions simply disappear in the translation. Secondly, we argue that these questions can be formalized and that recent developments in mathematical logic, specifically in model theory, provide a technical apparatus to discuss them.

Whitehead and Russell demonstrated that all of mathematics could be developed from a basis of logic and set theory. We are not concerned here with distinguishing between logic and set theory, or with ontological concerns about these basic propositions. Rather, we want to consider the difficulties with the underlying assumption that every mathematical proof can be translated into a formal argument from a few basic propositons via certain acceptable means of argument. We do not deny the truth of this assertion but we do deny the usual implicit conclusion that mathematical philosophy is thereby reduced to the study of which basic propositions and rules of inference should be allowed. Note that while Brouwer presumably intended otherwise intuitionism fits into this rubric. That is, the major difference between a formalist and classical mathematics is the choice of allowable rules of inference.

Tymoczko [] has argued that a mathematical proof should be convincing, surveyable, and formalizable. Note in passing that these are very nearly contradictory requirements. A fully formalized proof is neither convincing nor surveyable. Thus, at the very least Tymoczko demands a metamathematical component of every proof - a surveyable argument showing that what has been previously asserted can be formalized.

What does it mean to reduce mathematics to certain fundamental propositions? What properties are preserved by the reduction? Truth is preserved. Indeed, an informal argument yields a true consequence if and only if the translated version does. But, meaning is not preserved. We are all familiar with the fact that a mathematical argument can be clearly and convincingly presented in a few lines. But, crossing the t's and dotting the i's takes twenty pages and is incomprehensible.

This dichotomy is illustrated again by considering 'computer proof'. Does Hao Wang's computer generated proof of a theorem of the propositional caluculus convince. Can it? Does a formula of the propositional caluculus with more than twenty symbols have a meaning?

Here is a more precise example of the information lost by a truth-preserving transformation. Gurevich and Shelah have proved that it consistent that second order logic can be interpreted (in the usual syntactic sense) into the monadic theory of order. However, there is no first order interpretation [see Baldwin-Shelah]. That is, there is no definition of quantification over arbitrary relations in the monadic theory of order despite the fact that the two theories (consistently) have the same Turing degree.

The inuitive notion of ‘equivalence’ between proofs is also lost by this reduction. Certainly all proofs of the same result have become the same. I intend to argue that model theory has provided some formal tools for recognizing the similarity of proofs.

2 What are we talking about?

‘By an abstract entity I shall mean an entity that is neither spatial nor temporal and is neither a product of nor a process in a mind.’ [Resnik page 26] It is apparent to me that the objects of mathematics, whether something complicated like a spectral sequence or something more conceivable like a measurable cardinal are abstract entities. I do not however wish to commit myself to (or even think about) the ontology of such objects. Some may want to engage in some elaborate reductionism to preserve a more privileged state of reality for sense experience on the one hand or intuition on the other. While these are important questions in the philosophy of mathematics they are not our current subject. Similary, we ignore the question of how these abstract entitities are apprehended. We simply note that mathematicians are able to discuss any number of abstract objects without any real problem of subjectivism.

Note that from the individual mathematicians standpoint the complexity (abstractness) of an object depends on the number of levels of abstraction involved in defining it, not on its size. It is absurd to regard the third commutator group of a Borel subgroup of $SL(715, 11)$ as more concrete than $\aleph_{17}$ simply because the first is finite and the second infinite.

Mathematicians develop through study and conversation an intuition about certain abstract objects. They (discover or invent) formal descriptions of these objects. Depending upon the richness of these descriptions interesting results may or not be provable about the object. The description ‘17th infinite cardinal’ describes a certain object almost completely and very little of interest can be discovered about this object. Some interesting results appear when one considers the continuum - an object about which our current formal description is extremely incomplete. One can then show the independence of the assertion $2^{\aleph_0} < \aleph_{17}$. But everyone knows this is a theorem about $2^{\aleph_0}$ not about $\aleph_{17}$.

3 The body of the argument

(first try) (February 4, 1987)

Our initial argument is by example. We will consider several important mathematical theorems and remark informally on the key ideas of the proofs. We then note that these ideas have been formalized in model theory and thereby provide a common explanation of what at first appear to be only analogies. As Kreisel pointed out in considering the foundations of abstract recursion theory, in order to obtain the proper generalizations from theorems about finite or countable objects to theorems about objects of arbitrary cardinality it is necessary to clarify the meaning of finite.

We want to consider the relationship between two concepts: semisimple and completely reducible. To say that a group (ring) is semisimple is to say that a certain normal subgroup (ideal) is trivial. An algebra is completely reducible if it is a direct product of simple algebras. Now the prototheorem we want to examine asserts: every semisimple algebra is completely reducible.

The most well-known exemplar of this theorem is the Wedderburn-Artin theorem. Every semisimple ring with the descending chain condition on left ideals is completely reducible. Since the descending chain condition treats all left ideals the hypothesis of this theorem is phrased in second order rather than first order logic. However, it is a consequence of a more general theorem which can be phrased entirely in first order (albeit metamathematical) terms: every stable semisimple ring is completely reducible. To understand this result we need to know the definition of a stable structure. A structure (e.g. ring or group) G is stable if for every first order formula $\phi(\overline{x}, \overline{y})$ there is an integer $n(\phi)$ such that if a sequence $\langle \overline{a}_i : i < m \rangle$ of finite sequences of element of G is linearly ordered by ϕ then $m \leq n(\phi)$. In particular, in the ring theoretic case this implies the descending chain condition on principal left ideals and the ascending chain condition on principal annihilators. These are the only consequences of dcc used in the proof of the Wedderburn-Artin theorem.

We now consider some group theoretic versions of our featured proposition. Fitting [1938] proved that every finite semisimple group G contains a unique maximal centerless completely reducible (ccr) subgroup $\hat{G}$ and that $C_G(\hat{G})$, the set of elements of G which commute with each member of $\hat{G}$ is trivial. Thus, G can be faithfully embedded into $\mathrm{Aut}(\hat{G})$ (the automorphism group of $\hat{G}$) by conjugation. The finiteness of G is used in two distinct ways.

We are able to take a maximal subgroup satisfying a certain condition and we can induct on the cardinality of G.

The most successful versions of this theorem for infinite groups are for the algebraic groups. A crucial ingredient in the successful analysis of algebraic groups is the restriction of attention to definable subgroups. The algebraic group theorist calls a group, 'closed in the Zariski topology' if it is the collection of solutions of a set of polynomial equations. The fact that a polynomial ring over a field is Noetherian yields to a descending chain condition on closed subgroups. The replacement for the finite cardinality of the group is its dimension as an algebraic variety. By introducing one further notion, we are able to obtain an ascending chain condition. A group is connected if it has no definable (i.e. closed) subgroups of finite index. Now, if one connected group is properly contained in another the second must have larger dimension. Every algebraic group has finite dimension and thus it satisfies the ascending chain condition on connected subgroups. Repeating the arguments from the finite case, we can show every connected semisimple algebraic group can be embedded in the group of automorphisms of a completely reducible group. In fact, we can restrict the range to the algebraic (or definable) automorphisms. Now the theorem can be strengthened. One can show that if $\hat{G}$ is connected and semisimple (in particular ccr) then $\hat{G}$ has finite index in its group of definable automorphisms. Thus we obtain: if G is connected semisimple algebraic group then G is a direct product of simple groups.

(Aside: The usual statement is slight more complicated. To treat a number of interesting groups the definition of simple is often modified in the category of algebraic groups to say G is simple if G mod its center is simple in the usual sense. For our purposes this extension can be ignored; since it complicates statements and proofs without presenting any really new ideas.)

The notion of algebraic group might itself benefit from the model theorists analysis. On the other hand the following alternative might seem simpler only to a model theorist. Originally, an algebraic group is defined representationally. The group G is algebraic if it is isomorphic to a closed (in the Zariski topology) subgroup of a matrix group. In fact, however, Weil sought to isolate many properties of the group (or indeed of any algebraic variety) which arise from the existence of an embedding in affine or projective space but which do not depend upon the particular embedding. He also wanted to find a common rubric to describe both projective and affine varieties. For this purpose he invented the notion of an abstract algebraic variety. A complete description of this notion requires the introduction of a certain sheaf of functions and the definition of an algebraic morphism. An algebraic group is then an abstract algebraic variety with a binary and a unary operation which satisfy the axioms of groups and are morphisms of algebraic varieties. Hrushovski has shown that a group is algebraic over an algebraically closed field k. if and only if it can be defined in a quotient of k^n for some n. Then one can describe an algebraic group in a purely first order way. An algebraic group G is a structure whose universe is some subset of $k^n/\approx$ and whose relations are all those definable in the field k. I only suggest that one compare the length of the two definitions here and reflect that much more is left unsaid in the algebraic version.

There are several model theoretic counterparts to the representation of a semisimple algebraic group as a direct product of simple groups. These arise essentially from different generalizations of the finiteness conditions described above. Stability theory divides all structures into four classes, unstable on the one hand and on the other in increasing strength, stable, superstable, and ω-stable. We can approximate the meanings of the last two classes by the following facts. An ω-stable group satisfies the descending chain condition on definable subgroups. A superstable group has no descending chain of definable subgroups each with infinite index in its predecessor. Each superstable or ω stable group can be assigned an ordinal rank. An algebraic group over an algebraically closed field is ω-stable with finite rank. A structure A is $\aleph_1$-categorical if up to isomorphism there is only one structure of power $\aleph_1$ which satisfies the same first order sentences as A. Every simple algebraic group over an algebraically closed field is $\aleph_1$-categorical.

Now, Lascar has proved that every semisimple ω-stable group of finite rank is a direct product of $\aleph_1$-categorical groups. Baldwin and Pillay have shown that every semisimple superstable group is in the definable closure of a definable ccr subgroup $\hat{G}$. That is, for every $g \in G$, there is a formula $\phi(x, \overline{a})$ with parameters in $\hat{G}$ which is satisfied only by g. Both proofs use the chain conditions and induction on rank as a substitute for cardinality.

In this example, the tools of model theory have provided formal means to explicate two uses of finiteness, induction on cardinality and the existence of maximal subgroups satisfying certain conditions.