I

Truth and necessity in mathematics*

I hope that no one will think that there is no connection between the philosophy of the formal sciences and the philosophy of the empirical sciences. The philosophy that had such a great influence upon the empirical sciences in the last thirty years, the so-called 'logical empiricism' of Carnap and his school, was based upon two main principles:

(1) That the traditional questions of philosophy are so-called 'pseudo-questions' (*Scheinprobleme*), i.e. that they are wholly senseless; and

(2) That the theorems of the formal sciences – logic and mathematics – are analytic, not exactly in the Kantian sense, but in the sense that they 'say nothing', and only express our linguistic rules.

Today analytical philosophers are beginning to construct a new philosophy of science, one that also wishes to be unmetaphysical, but that cannot accept the main principles of 'logical empiricism'. The confrontation with the positivistic conception of mathematics is thus no purely technical matter, but has the greatest importance for the whole conception of philosophy of science.

What distinguishes statements which are true for mathematical reasons, or statements whose falsity is mathematically impossible (whether in the vocabulary of 'pure' mathematics, or not), or statements which are mathematically necessary, from other truths? Contrary to a good deal of received opinion, I propose to argue that the answer is *not* 'ontology', *not* vocabulary, indeed nothing 'linguistic' in any *reasonable* sense of linguistic. My strategy will not be to offer a contrary thesis, but rather to call attention to facts to which 'ontological' accounts and 'linguistic' accounts do not do justice. In the process, I hope to indicate just how complex are the facts of mathematical life, in contrast to the stereotypes that we have so often been given by philosophers as well as by mathematicians pontificating on the nature of their subject.

The idea that the 'ontology' (i.e. the domain of the bound variables) in a mathematically true statement is a domain of sets or numbers or

* A German version of this paper, titled 'Wahrheit und Notwendigkeit in der Mathematik', was presented as a public lecture at the University of Vienna on 3 June 1964, under the auspices of the Institut für Höhere Studien und Wissenschaftliche Forschung (Ford Institute).

functions or other 'mathematical objects', and (moreover) that *this* is what distinguishes mathematics from other sciences is a widespread one. On this view, mathematics is distinguished from other sciences by its *objects*, just as botany is distinguished from marine biology by the difference in the objects studied. This idea lives on in a constant tension with the other idea, familiar since Frege and Russell, that there is no sharp separation to be made between *logic* and mathematics. Yet logic, as such, has no 'ontology'! It is precisely the chief characteristic of the principles and inference rules of logic that *any* domain of objects may be selected, and that any expressions may be instantiated for the predicate letters and sentential letters that they contain. (I do not here count set theory, or higher order quantification theory, as 'logic'.) *Prima facie*, then, there is something problematical about one or both of these views.

In point of fact, it is not difficult to find mathematically true statements which quantify only over material objects, or over sensations, or over days of the week, or over whatever 'objects' you like: mathematically true statements about Turing machines, about inscriptions, about maps, etc. Thus, let T be a physically realized Turing machine, and let $P_1, P_2, \ldots, P_n$ be predicates in ordinary 'thing language' which describe its states. (E.g. P_1 might be: 'ratchet G1 of T is pressing against bar T4, etc.') An atomic instruction might be: 'If $P_2(T)$ and T is scanning the letter "y", T will erase the "y", print "z" in its stead, shift one square left on the tape (more tape will be adjoined if T ever reaches the end), and then adjust itself so that $P_6(T)$'. Such an instruction is wholly in 'nominalistic language' (does not quantify over abstract 'entities'). T is completely characterized by a finite set of such instructions, $I_1, I_2, \ldots, I_k$. Then the statement

(1) As long as I_1 and I_2 and ... and I_k, then T does not halt.

could very well be a mathematically true statement, and quantifies only over physical objects.

Again, suppose we use the symbols 1, 11, 111, ... to designate the numbers one, two, three... (i.e. the name of the number n is a string of n '1's'). In this notation, the sum of two numbers can be obtained by merely concatenating the numerals: nm is always the sum of n and m. Let us write $x = y^*$ to mean 'x equals y cubed', Nx to mean 'x is a number', and ! (read: 'shriek') to indicate absurdity. The following is a rather rudimentary formal system whose axioms can be seen to be correct on this interpretation:

System *E.S.*

Alphabet 1, ., =, *, !, N

Axioms

$$N1$$
$$Nx \rightarrow Nx1$$
$$Nx \rightarrow x = x$$
$$Nx \rightarrow x.1 = x$$
$$x.y = z \rightarrow x.y1 = zx$$
$$x.x = y, x.y = z \rightarrow z = x^*$$
$$z_1 = x_1^*, z_2 = x_2^*, z_3 = x_3^*, z_1 = z_2z_3 \rightarrow !$$

It is easily seen that ! is a theorem of *E.S.* if and only if some cube is the sum of two cubes. Fermat proved that this is impossible. Thus the following is true:

(2) If X is any finite sequence of inscriptions in the alphabet 1, ., =, *, !, N and each member of X is either an inscription of $N1$, or of a substitution instance of one of the remaining above axioms, or comes from two preceding terms in the sequence by Detachment, then X does *not* contain !.

Now, a finite *sequence* of inscriptions $I_1, \ldots, I_n$ can itself be identified with an inscription – say, with $I_1\#I_2\#\ldots\#I_n$, where # is a symbol not in the alphabet 1, ., =, *, !, N, which we employ as a 'spacer'. (We say that such an inscription has an inscription I in the alphabet 1, ., =, *, ! as a member, or 'contains' I, just in case the string begins with $I\#$, or ends with $\#I$, or has a proper part of the form $\#I\#$.) Thus (2) is once again an example of a mathematically true statement which refers only to physical objects (inscriptions). And such examples could easily be multiplied. Thus we see that, even if no one has yet succeeded in translating *all* of mathematics into 'nominalistic language', still there is no difficulty in expressing a *part* (and, indeed, a significant part) of mathematics in 'nominalistic language'.

Let me now, in the fashion of philosophers, consider an Objection to what I have said.

The Objection is that, even if some mathematically true statements quantify only over physical objects, still the *proofs* of these statements would refer at least to *numbers*, and hence to 'mathematical objects'. The reply to the Objection is that the premise is false! The principle needed to prove (2), for example, is the principle of Mathematical Induction. And this can be stated directly for finite inscriptions, and, moreover, can be perceived to be evidently true when so stated. It is not that one *must* state the principle first for numbers and *derive* the principle for inscriptions *via* goedel numbering. (Indeed, this would assume that every inscription possesses a goedel number, which cannot be proved without assuming the principle for inscriptions.) Rather, the principle can be seen to be evidently correct for arbitrary finite inscriptions, just as one sees intuitively that the principle of Induction in number theory is correct.

Here is the Principle, state as a Rule of proof:

$$\frac{\begin{array}{l} P(1) \\ P(.) \\ P(=) \\ P(*) \\ P(!) \\ P(N) \\ P(x) \rightarrow P(x1) \\ P(x) \rightarrow P(x.) \\ \quad\cdots\cdots \\ P(x) \rightarrow P(xN) \end{array}}{P(x)}$$

(In words: If 1, ., =, *, !, N are all P, and if, for every x, if $P(x)$ then $P(x1)$, $P(x.)$, . . ., $P(xN)$, then, for every x, $P(x)$.)

It would (I assert) be easy to write down a set of axioms about finite inscriptions (*not* including any axiom of infinity, since we are interested only in proving universal statements, and many of these do not require an assumption concerning the *existence* of inscriptions in their proof), from which by the Rule of Induction just stated and first order logic one could prove (2), and many similar statements besides. Moreover, if one is not content to simply *assume* the Rule of Induction, but wishes to 'derive' it, one can do *that* too! It suffices to assume Goodman's calculus of individuals (parts and wholes), instead of the theory of sets (or higher order quantification theory), as was done by Frege and Russell.† Thus the Objection falls, and I conclude that, whatever may be *essential* to mathematics, reference to abstract 'entities' is *not*. Indeed, one could even imagine a culture in which this portion of mathematics – parts and wholes, inscriptions in a finite alphabet, etc. – might be brought to a

† Quine has shown (Quine, 1964) that the Frege–Russell 'derivation' of mathematical induction can be redone so as to require only finite sets. Quine's methods can also be used, to the same effect, in the Theory of Inscriptions, with finite *wholes* playing the role of *sets*. It should be noted that the principle of induction that Russell claimed to be analytic – that every hereditary property which is possessed by zero is possessed by every number – is *empty* without further 'comprehension axioms' asserting that specific conditions define 'properties' (or sets). What plays the role of an *Aussonderungsaxiom* in the theory of finite wholes is the statement that 'there is a whole, *y*, which is got by joining all those *P*'s which are parts of *x*', where *x* is any whole. If we assume this as an axiom schema, then we can derive the rule of mathematical induction for finite inscriptions in the form given in the text from a definition of 'finite inscription' which is analogous to Quine's definition of number.

high degree of development without anyone ever having mentioned a so-called 'abstract entity'. What seems to characterize mathematics is a certain *style of reasoning*; but that style of reasoning is not essentially connected with an 'ontology'.

The same examples also go to show that, whatever it may be that distinguishes mathematically true statements from other statements, it is not *vocabulary*. Is it anything 'linguistic'? Consider the following statement:

(3) No one will ever draw a (plane) map which requires five or more colors to color. (The restriction being understood that two adjacent regions may not be colored the same.)

Do the Rules of English, in any reasonable sense of 'rule of English', decide whether (3) is the sort of statement that is true (or false) for mathematical rather than for empirical reasons? If we take the 'rules of English' to be the rules of 'generative grammars', such as the ones constructed by Noam Chomsky and his followers, or even the State Regularities and Semantic Rules espoused (respectively) by Paul Ziff and by Katz and Fodor (see the references at the end of this paper), then the answer is clearly 'no'. Some philosophers write as if, in addition to these everyday or 'garden variety' rules of English, which are capable of being discovered by responsible linguistic investigation carried on by trained students of language, there were also 'depth rules' capable of being discovered only by *philosophers*, and as if these in some strange way *accounted for* mathematical truth. Although this is a currently fashionable line, I am inclined to think that all this is a mare's nest. But without allowing myself to be drawn into a discussion of this 'depth grammatical' approach to mathematical truth at this point, there is a point which I would still call to your attention: namely, if one says that the 'rules of the language' *decide* whether (3) is a mathematically true or empirically true (or false) statement, then they must *also* decide the mathematical question *can every map be colored with four colors*? For suppose that (3) is mathematically true, i.e. true because what it says will never be done *could* never be done (in the mathematical sense of 'could'). Then (3) would be analogous to:

No one will every exhibit a cube which is the sum of two cubes.

In this case the answer to the Four Color Problem is 'yes' – every (plane) map *can* be colored with four colors. On the other hand, suppose (3) is an *empirical* statement. Then it must be that it is *possible* (mathematically) to produce a map which requires five or more colors to color. For this is the only case in which it *is* an empirical question whether or not anyone ever *will* produce a map which requires five or more colors to color. Thus, any system of Linguistic Rules which is powerful enough to decide whether (3) is an empirical statement or a mathematically necessary statement will also have to be powerful enough to decide the Four Color Problem. And, in general, if the distinction between statements whose truth (or falsity) is mathematical in character and statements whose truth (or falsity) is empirical in character is drawn by the Rules of the Language, then all mathematical truth (in combinatorial mathematics, anyway) is decided by the Rules of the Language. The apparently modest thesis that purely linguistic considerations can determine what is a question of mathematical truth or falsity and what is a question of empirical (or ethical, or theological, or legal, or anyway other-than-mathematical) truth or falsity, is in reality no more modest than the radical thesis that linguistic considerations can determine not just that 'is it the case that *p*?' is a mathematical or empirical or whatever question, but also can determine the *answer* to the question, when it is a mathematical one.

Not only do 'linguistic' considerations, in any reasonable sense, seem unable to determine what is a mathematically necessary (or impossible) statement, and what is a contingent statement, but it is doubtful if they can single out even the class of assertions of *pure* mathematics. (Consider assertions like (2), for example, some of which belong to pure mathematics, on any reasonable definition, while some – of the same general form as (2), and referring to systems like *E.S.* – are only contingently true, i.e. true because some mathematically possible finite inscriptions in the alphabet in question do not in fact exist.) However, this difficulty can be overcome by 'cheating': we can specify that an assertion will not *count* as a statement of 'pure' mathematics unless there is some indication either in the wording (e.g. the presence of such an expression as 'mathematically implies'), or in the context (e.g. the assertion appears as the last line of a putative mathematical proof) that the statement is supposed to be true for mathematical reasons.

Once again, I shall succumb to my professional habit of considering Objections. The Objection this time is that (3) has (allegedly) two different *senses*: that (3) means one thing when it is intended as a mathematical statement, and something different when it is intended as an empirical statement. The Reply to the Objection is: what if I simply intend to say that no one will ever do the thing specified? Is it really to be supposed that this statement is *ambiguous*, and that the hearer does not know what I *mean* unless he knows my *grounds*? I believe that Wittgenstein would say that the ambiguity is as follows: if I intend (3) as a *prediction*, then I am not *using* (3) as a mathematical statement; but if I am not prepared to count anything as a counter-example to (3), then I am using (3) as a mathematical statement. It will be the burden of the remainder of this paper to criticize this account of what it is to accept a statement as mathematically necessary.

The revisability of mathematical assertions

According to a currently fashionable view, to accept a statement as mathematically true is in some never-quite-clearly-explained way a matter of accepting a rule of language, or a 'stipulation', or a 'rule of description', etc. Of a certainty, the concepts of possibility and impossibility are of great importance in connection with mathematics. To say that a statement is mathematically true is to say that the negation of the statement is mathematically impossible. To say that a statement is mathematically true is to say that the statement is mathematically necessary. But I cannot agree that Necessity is the same thing as Unrevisability. No one would be so misguided as to urge that 'mathematically necessary' and 'immune from revision' are synonymous expressions; however, it is another part of the current philosophical fashion to say 'the question is not what does X mean, but what does the *act of saying X* signify'; and I fear that the view may be widespread that the act of saying that a proposition p is mathematically necessary, or even of asserting p on clearly mathematical grounds, or in a clearly mathematical context, *signifies* that the speaker would not count anything as a counter-example to p, or even adopts the rule that nothing *is to count* as a counter-example to p. This view is a late refinement of the view, which appears in the writings of Ayer and Carnap, that mathematical statements are consequences of 'semantical rules'. The Ayer–Carnap position is open to the crushing rejoinder that being a *consequence* of a semantical rule is *following mathematically* from a semantical rule; so all that has really been said is that *mathematics* = language plus *mathematics*. This older view went along with the denial that there is any such thing as Discovery in mathematics; or, rather, with the assertion that there is discovery 'only in a psychological sense'. The information given by the conclusion is already contained in the premisses, in a mathematically valid argument, these philosophers said: only we are sometimes 'psychologically' unable to see that it is. On examination, however, it soon appeared that these philosophers were simply *redefining* 'information' so that mathematically equivalent propositions are *said* to give the same information. It is only in this Pickwickian sense of 'information' that the information given in the conclusion of a deductive argument is already given in the premisses; and our inability to see that it is requires no special explanation at all, 'psychological' or otherwise. On the contrary, what *would* require psychological explanation would be the *ability*, if any creature ever possessed it, to infallibly and instantly see that two propositions were mathematically equivalent, regardless of differences in syntactic structure and in semantic interpretation in the *linguistic* sense of 'semantic interpretation'. Far from showing that Discovery is absent in mathematics, Carnap and Ayer only enunciated the uninteresting tautology that in a valid mathematical argument the conclusion is mathematically implied by the premisses.

The late refinement described above does avoid the emptiness of the Ayer–Carnap view. To accept a statement as mathematically necessary is then and there to adopt a 'rule'; not to recognize that in some sense which remains to be explained the statement is a *consequence* of a rule. Thus the late refinement is both more radical and more interesting than the older view. We adopt such a rule not *arbitrarily*, but because of our Nature (and the nature of the objects with which we come into contact, including Proofs). That a sequence of symbols exists which impels us to say 'it is mathematically impossible that a cube should be the sum of two cubes' is indeed a *discovery*. But we should not be misled by the picture of the Mathematical Fact waiting to be discovered in a Platonic Heaven, or think that a given object would be a proof if our nature did not impel us to *employ* it as a proof.

What student of the philosophy of mathematics has not encountered this sort of talk lately? And what *intelligent* student has not been irritated by it? *To be sure*, a proof is more than a mere sequence of marks, on paper or in the sand. If the symbols did not have meaning in a language, if human beings did not speak, did not do mathematics, did not follow proofs and so on, the same sequence of marks would *not* be a proof. But it would still be true that no cube is the sum of two cubes (even if no one proved it). Indeed, these philosophers do not deny this later assertion; but they say, to repeat, that we must not be misled by the picture of the Mathematical Fact as Eternally True, independent of Human Nature and human mathematical activity. But I, for one, do not find this picture at all misleading. Nor have I ever been told just *how* and *why* it is supposed to be misleading. On the contrary, I allege that it is the picture according to which accepting a mathematical proposition is accepting a 'rule of description' that is radically misleading, and in a way that is for once definitely specifiable; *this* picture is misleading in that it suggests that once a mathematical assertion has been accepted by me, I will not allow anything to *count* against this assertion. This is an *obviously* silly suggestion; but what is left of the sophisticated view that we have been discussing once this silly suggestion has been repudiated?

Indeed, not only can we change our minds concerning the truth or falsity of mathematical assertions, but there are, I believe, different *degrees* of revisability to be discerned even in elementary number theory.

To illustrate this last contention, consider the statement 'no cube is the sum of two cubes'. I accept this statement; I know that it can be proved in First Order Arithmetic. Now, let us suppose that I cube 1769 and add the cube of 269. I discover (let us say) that the resulting number is exactly the cube of 1872. Is it really supposed that I would not allow this empirical event (the calculation, etc., that I do) to *count against* my mathematical assertion? To accept *any* assertion p is to refuse to accept $\bar{p}$ *as long as I still accept p*. (The Principle of Contradiction.) So, to accept $x^3+y^3 \neq z^3$ is to refuse to accept the calculation *as long as I still accept* $x^3+y^3 \neq z^3$. But there is all the difference in the world between this and adopting the convention that no experience in the world is to be *allowed to refute* $x^3+y^3 \neq z^3$ – which would be insane.

But what would I in fact do if I discovered that $(1769)^3+(269)^3 = (1872)^3$? I would first go back and look for a mistake in the proof of $x^3+y^3 \neq z^3$. If I decided that the latter proof was a correct proof in First Order Arithmetic, then I would have to modify First Order Arithmetic – which would be shattering. But it is clear that in such revision purely *singular* statements ($5+7 = 12$, $189+12 = 201$, $34 \cdot 2 = 68$, etc.) would take priority over *generalized* statements ($x^3+y^3 \neq z^3$). A generalized statement can be refuted by a singular statement which is a counter-example.

Gentzen has given a convincing example of a constructive proof (of the consistency of First Order Arithmetic, in fact), which is not formalizable in First Order Arithmetic. This proof has been widely misunderstood because of its use of transfinite induction. The point is not that Gentzen used transfinite methods, but that, on the contrary, transfinite methods can be *justified* constructively in some cases. For ε_0 induction is not taken as primitive in constructive mathematics. On the contrary, for $\alpha < \varepsilon_0$, α-induction can be reduced to ordinary induction, even in Intuitionistic Arithmetic. And if we adjoin to Intuitionistic Arithmetic the constructively self-evident principle: if *it is provable in the system* (Intuitionistic Arithmetic) *that for each numeral n*, $F(n)$ *is provable, then* $(x)F(x)$ – then we can formalize Gentzen's whole proof, even though our bound variables range only over numbers, and not over 'ordinals' (in the classical sense), at all.

I point this out, because I believe that it is important that there are methods of proof which are regarded as completely secure by everyone who understands them, and which are wholly constructive, but which go outside First Order Arithmetic. Now, suppose that we had such a proof that established the ω-inconsistency of First Order Arithmetic. (In fact, this cannot happen – Charles Parsons has given constructive proof of the ω-consistency of First Order Arithmetic.) Imagine that the proof is entirely perspicuous, and that after reading it we are convinced that we could actually construct proofs of $\sim F(0)$, $\sim F(1), \ldots$, for some formula F such that $(Ex)F(x)$ has been proved in First Order Arithmetic. If F is a decidable property of integers, this amounts to saying that $F(0)$, $F(1)$, $F(2), \ldots$ are all *false* even though $(Ex)Fx$ is *provable*. And *this* amounts to saying that First Order Arithmetic is *incorrect*; since $(Ex)Fx$, although *provable* is clearly *false*.

I see nothing unimaginable about the situation just described. Thus, just as accepting $x^3+y^3 \neq z^3$ does not incapacitate me from recognizing a counter-example if I ever come across one ($(1769)^3+(269)^3 = (1872)^3$ would be a counter-example if the arithmetic checked): so accepting $(Ex)Fx$ does not incapacitate me from recognizing a disproof if I see one (such a disproof would establish, *by means more secure than those used in the proof of* $(Ex)Fx$, that $\sim F(0)$, $\sim F(1)$, $\sim F(2), \ldots$ would all prove correct on calculation). And, in general, even if a mathematical statement has been *proved*, its falsity may not be literally impossible to imagine; and there is no rule that 'nothing is to count against this'.

Another kind of revision that takes place in mathematics is too obvious to require much attention: we may *discover a mistake in the proof*. It is often suggested that this is not *really* changing our minds about a *mathematical* question, but why not? The question, 'is this object in front of me a proof of S in *E.S.*'? may not be a mathematical question, but am I not *also* changing my mind about the question 'is it true that, S?'? And this latter question *may* be a mathematical question. Indeed, if I previously adopted the stipulation that *nothing was to count against* S, then *how could* my attitude towards the mathematical proposition S be affected by the brutally empirical discovery that a particular sequence of inscriptions did not have the properties I thought it had?

An especially interesting case arises when S is itself a proposition concerning proof in a formal system, say 'F is provable in L'. Such facts are especially recalcitrant to the 'rule of description' account, as are indeed all purely existential combinatorial facts, – e.g. 'there exists an odd perfect number', if that statement is true. We may change our minds about such a statement indefinitely often (if the proof is very long); our attitude towards it depends upon brutally empirical questions of the kind mentioned in the preceding paragraph; there may be no *reason* for such a fact apart from the fact that there just *is* a proof of F in L – i.e. we often have no way of establishing such a statement except