

THE FOUNDATIONS OF MATHEMATICS

By F. P. RAMSEY.

[Received 23 August, 1925.—Read 12 November, 1925.]

Preface.

THE object of this paper is to give a satisfactory account of the Foundations of Mathematics, in accordance with the general method of Frege, Whitehead, and Russell. Following these authorities, I hold that mathematics is part of logic, and so belong to what may be called the logical school as opposed to the formalist and intuitionist schools. I have therefore taken *Principia Mathematica* as a basis for discussion and amendment, and believe myself to have discovered how, by using the work of Mr. Ludwig Wittgenstein, it can be rendered free from the serious objections which have caused its rejection by the majority of German authorities who have deserted altogether its line of approach.

Contents.

- I. Introduction.
- II. *Principia Mathematica*.
- III. Predicative Functions.
- IV. Functions in Extension.
- V. The Axioms.

I. *Introduction.*

In this chapter we shall be concerned with the general nature of pure mathematics*, and how it is distinguished from other sciences.

* In future by "mathematics" will always be meant "pure mathematics."

Here there are really two distinct categories of things of which an account must be given, the ideas or concepts of mathematics, and the propositions of mathematics. This distinction is neither artificial nor unnecessary, for the great majority of writers on the subject have concentrated their attention on the explanation of one or other of these categories, and erroneously supposed that a satisfactory explanation of the other would immediately follow.

Thus the formalist school, of whom the most eminent representative is now Hilbert, have concentrated on the propositions of mathematics, such as " $2+2=4$ ", which they have pronounced to be meaningless formulae to be manipulated according to certain arbitrary rules, and mathematical knowledge they hold to consist in knowing what formulae can be derived from what others consistently with the rules. Such being the propositions of mathematics, the account of its concepts, for example the number 2, immediately follows. "2" is a meaningless mark occurring in these meaningless formulae. But, whatever may be thought of this as an account of mathematical propositions, it is obviously hopeless as a theory of mathematical concepts; for these occur not only in mathematical propositions, but also in those of everyday life. Thus "2" occurs not merely in " $2+2=4$ ", but also in "It is 2 miles to the station", which is not a meaningless formula, but a significant proposition, in which "2" cannot conceivably be a meaningless mark. Nor can there be any doubt that "2" is used in the same sense in the two cases, for we can use " $2+2=4$ " to infer from "It is two miles to the station and two miles on to the Gogs", that "it is four miles to the Gogs via the station", so that these ordinary meanings of two and four are clearly involved in " $2+2=4$ ". So the hopelessly inadequate formalist theory is, to some extent, the result of considering only the propositions of mathematics, and neglecting the analysis of its concepts, on which additional light can be thrown by their occurrence outside mathematics in the propositions of everyday life.

Apart from formalism, there are two main general attitudes to the foundation of mathematics: that of the intuitionists or finitists like Brouwer and Weyl in his recent papers, and that of the logicians, Frege, Whitehead, and Russell. The theories of the intuitionists admittedly involve giving up many of the most fruitful methods of modern analysis, for no reason, as it seems to me, except that the methods fail to conform to their private prejudices. They do not, therefore, profess to give any foundation for mathematics as we know it, but only for a narrower body of truth which has not yet been clearly

defined. There remain the logicians whose work culminated in *Principia Mathematica*. The theories there put forward are generally rejected for reasons of detail, especially the apparently insuperable difficulties connected with the axiom of reducibility. But these defects in detail seem to me to be results of an important defect in principle, first pointed out by Mr. Wittgenstein.

The logical school has concentrated on the analysis of mathematical concepts, which it has shown to be definable in terms of a very small number of fundamental logical concepts, and, having given this account of the concepts of mathematics, they have immediately deduced an account of mathematical propositions, namely that they were those true propositions in which only mathematical or logical concepts occurred. Thus Russell, in the *Principles of Mathematics*, defines pure mathematics as "the class of all propositions of the form p implies q , where p and q are propositions containing one or more variables the same in the two propositions, and neither p nor q contains any constants except logical constants"* . This reduction of mathematics to symbolic logic was rightly described by Mr. Russell as one of the greatest discoveries of our age† ; but it was not the end of the matter, as he seemed to suppose, because he was still far from an adequate conception of the nature of symbolic logic, to which mathematics had been reduced. I am not referring to his naive theory that logical constants were names for real objects (which he has since abandoned), but to his belief that any proposition which could be stated by using logical terms‡ alone must be a proposition of logic or mathematics§ . I think the question is made clearer by describing the class of propositions in question as the completely general propositions, emphasizing the fact that they are not about particular things or relations, but some or all things and relations. It is really obvious that not all such propositions are propositions of mathematics or symbolic logic. Take, for example, "any two things differ in at least thirty ways"; this is a completely general proposition, it could be expressed as an implication involving only logical constants and variables, and it may well be true. But as a mathematical or logical truth no one could regard it; it is utterly different from such a proposition as "any two things together with any other two things make four things", which is a logical and not merely an empirical

* Russell, *Principles of Mathematics*, 1 (1903), 3.

† *Loc. cit.*, 5.

‡ *I.e.* variables and logical constants.

§ I neglect here, as elsewhere, the arbitrary and trivial proviso that the proposition must be of the form " p implies q ".

truth. According to our philosophy we may differ in calling the one a contingent, the other a necessary proposition, or the one a genuine proposition, the other a mere tautology; but we must all agree that there is some essential difference between the two, and that a definition of mathematical propositions must include not merely their complete generality but some further property as well. This is pointed out, with a reference to Wittgenstein, in Russell's *Introduction to Mathematical Philosophy**; but there is no trace of it in *Principia Mathematica*, nor does Mr. Russell seem to have understood its tremendous importance, for example, in the consideration of primitive propositions. In the passage referred to in the *Introduction to Mathematical Philosophy*, Mr. Russell distinguishes between propositions which can be enunciated in logical terms from those which logic can assert to be true, and gives as the additional characteristic of the latter that they are "tautological" in a sense which he cannot define. It is obvious that a definition of this characteristic is essential for a clear foundation of our subject, since the idea to be defined is one of the essential sides of mathematical propositions, their content, and their form. Their content must be completely generalized, and their form tautological.

The formalists neglected the content altogether and made mathematics meaningless, the logicians neglected the form and made mathematics consist of any true generalizations; only by taking account of both sides and regarding it as composed of tautologous generalizations can we obtain an adequate theory.

We have now to explain a definition of tautology which has been given by Mr. Wittgenstein in his *Tractatus Logico-Philosophicus*, and forms one of the most important of his contributions to the subject. In doing this we cannot avoid some explanation of his theory of propositions in general.

We must begin with the notion of an *atomic proposition*†; this is one which could not be analysed in terms of other propositions, and could consist of names alone without logical constants. For instance, by joining to " ϕ ", the name of a quality, " a ", the name of an individual, and writing " ϕa ", we have an atomic proposition, asserting that the individual has the quality. Thus, if we neglect the fact that "Socrates" and "wise" are incomplete symbols and regard them as names, "Socrates is wise" is an atomic proposition; but "All men are wise", "Socrates is not wise", are not atomic.

* Page 205.

† Wittgenstein calls these "elementary propositions"; I have called them "atomic" in order to follow Mr. Russell in using "elementary" with a different meaning.

Suppose now we have, say, n atomic propositions p, q, r, \dots . With regard to their truth or falsity there are 2^n mutually exclusive ultimate possibilities, which we could arrange in a table like this (T signifies truth, and F falsity, and we have taken $n = 2$ for brevity).

p	q
T	T
T	F
F	T
F	F

These 2^n possibilities we will call the truth-possibilities of the n atomic propositions. We may wish to pick out any sub-set of them, and assert that it is a possibility out of this sub-set which is, in fact, realized—that is, to express our agreement with some of the possibilities and our disagreement with the remainder. We can do this by setting marks T and F against the possibilities with which we agree and disagree respectively. In this way we obtain a proposition.

Thus

“

p	q	
T	T	F
T	F	T
F	T	T
F	F	T

”

is the proposition “not both p and q are true”, or “ p is incompatible with q ”, for we have allowed all the possibilities except the first, which we have disallowed.

Similarly

p	q	
T	T	T
F	T	T
T	F	F
F	F	T

is the proposition "if p , then q ".

A proposition which expresses agreement and disagreement with the truth-possibilities of p, q, \dots (which need not be atomic) is called a truth-function of the arguments p, q, \dots . Or, more accurately, P is said to be the same truth-function of p, q, \dots as R is of r, s, \dots if P expresses agreement with the truth-possibilities of p, q, \dots corresponding by the substitution of p for r, q for s, \dots to the truth-possibilities of r, s, \dots with which R expresses agreement. Thus " p and q " is the same truth-function of p, q as " r and s " is of r, s , in each case the only possibility allowed being that both the arguments are true. Mr. Wittgenstein has perceived that, if we accept this account of truth-functions as expressing agreement and disagreement with truth-possibilities, there is no reason why the arguments to a truth-function should not be infinite in number*. As no previous writer has considered truth-functions as capable of more than a finite number of arguments, this is a most important innovation. Of course, if the arguments are infinite in number they cannot all be enumerated and written down separately; but there is no need for us to enumerate them if we can determine them in any other way, as we can by using propositional functions.

A propositional function is an expression of the form " $f\hat{x}$ ", which is such that it expresses a proposition, when any symbol (of a certain appropriate logical type depending on f) is substituted for " \hat{x} ". Thus " \hat{x} is a man" is a propositional function. We can use propositional functions to collect together the range of propositions, which are all the values of the function for all possible values of x . Thus " \hat{x} is a man" collects together all the propositions " a is a man", " b is a man",

* Thus the logical sum of a set of propositions is the proposition that one at least of the set is true, and it is immaterial whether the set is finite or infinite. On the other hand, an infinite algebraic sum is not really a sum at all but a *limit*, and so cannot be treated as a sum except subject to certain restrictions.

etc. Having now by means of a propositional function defined a set of propositions, we can, by using an appropriate notation, assert the logical sum or product of this set. Thus, by writing " $(x) . fx$ ", we assert the logical product of all propositions of the form " fx "; by writing " $(\exists x) . fx$ " we assert their logical sum. Thus " $(x) . x$ is a man" would mean "everything is a man", " $(\exists x) . x$ is a man", "there is something which is a man". In the first case we allow only the possibility that all the propositions of the form " x is a man" are true; in the second, we exclude only the possibility that all the propositions of the form " x is a man" are false.

Thus general propositions containing "all" and "some" are found to be truth-functions, for which the arguments are not enumerated but given in another way. But we must guard here against a possible mistake. Take such a proposition as "all men are mortal"; this is not as might at first sight be supposed the logical product of the propositions " x is mortal" for such values of x as are men. Such an interpretation can easily be shown to be erroneous (see, for example, *Principia Mathematica*, 1st ed., I, 47). "All men are mortal" must be interpreted as meaning " $(x) .$ if x is a man, x is mortal", i.e. it is the logical product of all the values of the function "if x is a man, x is mortal".

Mr. Wittgenstein maintains that all propositions are, in the sense defined, truth-functions of elementary propositions. This is hard to prove, but is on its own merits extremely plausible; it says that, when we assert anything, we are saying that it is one out of a certain group of ultimate possibilities which is realized, not one out of the remaining possibilities. Also it applies to all the propositions which could be expressed in the symbolism of *Principia Mathematica*, since these are built up from atomic propositions by using, first, conjunctions like "if", "and", "or"; and, secondly, various kinds of generality (apparent variables). And both these methods of construction have been shown to create truth-functions*.

From this account we see when two propositional symbols are to be regarded as instances of the same proposition, namely, when they express agreement and disagreement with the same sets of truth-possibilities of atomic propositions.

Thus in the symbolism of *Principia Mathematica*

$$"p \supset q : \sim p . \supset . q", \quad "q \vee : p . \sim p"$$

are both more complicated ways of writing " q ".

* The form " A believes p " will perhaps be suggested as doubtful. This is clearly not a truth-function of " p ", but may nevertheless be one of other atomic propositions.

Given any set of n atomic propositions as arguments, there are 2^n corresponding truth-possibilities, and therefore 2^{2^n} sub-classes of their truth-possibilities, and so 2^{2^n} truth-functions of n arguments, one expressing agreement with each sub-class and disagreement with the remainder. But among these 2^{2^n} there are two extreme cases of great importance: one in which we express agreement with all the truth-possibilities, the other in which we express agreement with none of them. A proposition of the first kind is called a *tautology*, of the second a *contradiction*. Tautologies and contradictions are not real propositions, but degenerate cases. We may, perhaps, make this clear most easily by taking the simplest case, when there is only one argument.

The tautology is “ $\left. \begin{array}{l} p \\ T \quad T \\ F \quad T \end{array} \right\} ”, i.e. “p \text{ or not } p”$ ”.

This really asserts nothing whatever; it leaves you no wiser than it found you. You know nothing about the weather, if you know that it is either raining or not raining*.

The contradiction is “ $\left. \begin{array}{l} p \\ T \quad F \\ F \quad F \end{array} \right\} ”, i.e. “p \text{ is neither true nor false”$ ”.

This is clearly self-contradictory and does not represent a possible state of affairs whose existence could be asserted.

Tautologies and contradictions can be of all degrees of complexity; to give other examples “ $(x). \phi x : \supset : \phi a$ ” is a tautology. “ $\sim . (\exists x). \phi x : \phi a$ ” a contradiction. Clearly by negating a contradiction we get a tautology, and by negating a tautology a contradiction. It is important to see that tautologies are not simply true propositions, though for many purposes they can be treated as true propositions. A genuine proposition asserts something about reality, and it is true if reality is as it is asserted to be. But a tautology is a symbol constructed

* Wittgenstein, *Tractatus Logico-Philosophicus*, 4 · 461.

so as to say nothing whatever about reality, but to express total ignorance by agreeing with every possibility. The assimilation of tautologies and contradictions with true and false propositions respectively results from the fact that tautologies and contradictions can be taken as arguments to truth-functions just like ordinary propositions, and for determining the truth or falsity of the truth-function, tautologies and contradictions among its arguments must be counted as true and false respectively. Thus, if "t" be a tautology, "c" a contradiction, "t and p", "if t, then p". "c or p" are the same as "p", and "t or p". "if c, then p" are tautologies.

We have here, thanks to Mr. Wittgenstein, to whom the whole of this analysis is due, a clearly defined sense of tautology; but is this, it may be asked, the sense in which we found tautology to be an essential characteristic of the propositions of mathematics and symbolic logic? The question must be decided by comparison. Are the propositions of symbolic logic and mathematics tautologies in Mr. Wittgenstein's sense?

Let us begin by considering not the propositions of mathematics but those of *Principia Mathematica**. These are obtained by the process of deduction from certain primitive propositions, which fall into two groups; those expressed in symbols and those expressed in words. Those expressed in words are nearly all nonsense by the theory of types, and should be replaced by symbolic conventions. The real primitive propositions, those expressed in symbols, are, with one exception, tautologies in Wittgenstein's sense. So, as the process of deduction is such that from tautologies only tautologies follow, were it not for one blemish the whole structure would consist of tautologies. The blemish is, of course, the Axiom of Reducibility, which is, as will be shown below†, a genuine proposition, whose truth or falsity is a matter of brute fact, not of logic. It is, therefore, not a tautology in any sense, and its introduction into mathematics is inexcusable. But suppose it could be dispensed with, and *Principia Mathematica* were modified accordingly, it would consist entirely of tautologies in Wittgenstein's sense. And therefore if *Principia Mathematica* is on the right lines as a foundation and interpretation of mathematics, it is Wittgenstein's sense of tautology in which mathematics is tautologous.

But the adequacy of *Principia Mathematica* is a matter of detail, and, since we have seen it contains a very serious flaw, we can no longer be

* This is only because *Principia Mathematica* may be a wrong interpretation of mathematics; in the main I think it is a right one.

† See Chapter V.

sure that mathematics is the kind of thing Whitehead and Russell suppose it to be, or therefore that it consists of tautologies in Wittgenstein's sense. One thing is, however, clear: that mathematics does not consist of genuine propositions or assertions of fact which could be based on inductive evidence, as it was proposed to base the axiom of reducibility, but is in some sense necessary or tautologous. In actual life, as Wittgenstein says, "it is never a mathematical proposition which we really need, but we use mathematical propositions *only* in order to infer from propositions which do not belong to mathematics to others which equally do not belong to mathematics*". Thus we use " $2 \times 2 = 4$ " to infer from "I have two pennies in each of my two pockets" to "I have four pennies altogether in my pockets". " $2 \times 2 = 4$ " is not itself a genuine proposition in favour of which inductive evidence can be required, but a tautology which can be seen to be tautologous by anyone who can fully grasp its meaning. When we proceed further in mathematics the propositions become so complicated that we cannot see immediately that they are tautologous, and have to assure ourselves of this by deducing them from more obvious tautologies. The primitive propositions on which we fall back in the end must be such that no evidence could be required for them, since they are patent tautologies like "if p then p ". But the tautologies of which mathematics consist may perhaps turn out not to be of Wittgenstein's kind, but of some other. Their essential use is to facilitate logical inference; this is achieved in the most obvious way by constructing tautologies in Wittgenstein's sense, for if "if p , then q " is a tautology, we can logically infer " q " from " p ", and, conversely, if " q " follows logically from " p ", "if p , then q " is a tautology†. But it is possible that there are other kinds of formulae which could be used to facilitate inference; for instance, what we may call identities such as " $a = b$ ", signifying that " a ", " b " may be substituted for one another in any proposition without altering it. I do not mean without altering its truth or falsity, but without altering what proposition it is. " $2+2 = 4$ " might well be an identity in this sense, since "I have $2+2$ hats" and "I have 4 hats" are the same proposition, as they agree and disagree with the same sets of ultimate truth-possibilities.

Our next problem is to decide whether mathematics consists of tautologies (in the precise sense defined by Wittgenstein, to which we

* Wittgenstein, *op. cit.*, 6·211.

† This may perhaps be made clearer by remarking that if " q " follows logically from " p ", " $p \sim q$ " must be self-contradictory, therefore " $\sim(p \sim q)$ " tautologous or " $p \supset q$ " tautologous.

shall in future confine the word "tautology") or of formulae of some other sort. It is fairly clear that geometry, in which we regard such terms as "point", "line" as meaning any things satisfying certain axioms, so that the only constant terms are truth-functions like "or", "some", consists of tautologies. And the same would be true of analysis, if we regarded numbers as any things satisfying Peano's axioms. Such a view would, however, be certainly inadequate, because since the numbers from 100 on satisfy Peano's axioms, it would give us no means of distinguishing "This equation has three roots" from "This equation has a hundred and three roots." So numbers must be defined not as variables but as constants, and the nature of the propositions of analysis becomes doubtful.

I believe that they are tautologies, but the proof of this depends on giving a detailed analysis of them, and the disproof of any other theory would depend on finding an insuperable difficulty in the details of its construction. In this chapter I propose to discuss the question in a general way, which must inevitably be rather vague and unsatisfactory. I shall first try to explain the great difficulties which a theory of mathematics as tautologies must overcome, and then I shall try to explain why the alternative sort of theory suggested by these difficulties seems hopelessly impracticable. Then in the following chapters I shall return to the theory that mathematics consists of tautologies, discuss and partially reject the method for overcoming the difficulties given in *Principia Mathematica*, and construct an alternative and, to my mind, satisfactory solution.

Our first business is, then, the difficulties of the tautology theory. They spring from a fundamental characteristic of modern analysis which we have now to emphasize. This characteristic may be called *extensionality*, and the difficulties may be explained as those which confront us if we try to reduce a calculus of extensions to a calculus of truth-functions. Here, of course, we are using "extension" in its logical sense, in which the extension of a predicate is a class, that of a relation a class of ordered couples; so that in calling mathematics extensional we mean that it deals not with predicates but with classes, not with relations in the ordinary sense but with possible correlations, or "relations in extension," as Mr. Russell calls them. Let us take as examples of this point three fundamental mathematical concepts, the idea of a real number, the idea of a function (of a real variable), and the idea of similarity of classes (in Cantor's sense).

Real numbers are defined as segments of rationals; any segment of rationals is a real number, and there are 2^{\aleph_0} of them. It is not

necessary that the segment should be defined by any property or predicate of its members, in any ordinary sense of predicate. A real number is therefore an extension, and it may even be an extension with no corresponding intension. In the same way a function of a real variable is a relation in extension, which need not be given by any real relation or formula.

The point is perhaps most striking in the Cantor's definition of similarity. Two classes are said to be similar (*i.e.* have the same cardinal number) when there is a one-one relation whose domain is the one class and converse domain the other. Here it is essential that the one-one relation need only be a relation in extension; it is obvious that two classes could be similar—*i.e.* capable of being correlated without there being any relation actually correlating them.

There is a verbal point which requires mention here; I do not use the word "class" to imply a principle of classification, as the word naturally suggests, but by a "class" I mean any set of things of the same logical type. Such a set, it seems to me, may or may not be definable either by enumeration or as the extension of a predicate. If it is not so definable we cannot mention it by itself, but only deal with it by implication in propositions about all classes, or some classes. The same is true of relations in extension, by which I do not merely mean the extensions of actual relations, but any sets of ordered couples. That this is the notion occurring in mathematics seems to me absolutely clear from the last of the above examples, Cantor's definition of similarity, where obviously there is no need for the one-one relation in extension to be either finite or the extension of an actual relation.

Mathematics is therefore essentially extensional, and may be called a calculus of extensions, since its propositions assert relations between extensions. This, as we have said, is hard to reduce to a calculus of truth-functions, to which it must be reduced if mathematics is to consist of tautologies, for tautologies are truth-functions of a certain special sort, namely those agreeing with all the truth-possibilities of their arguments. We can perhaps most easily explain the difficulty by an example.

Let us take an extensional assertion of the simplest possible sort: the assertion that one class includes another. So long as the classes are defined as the classes of things having certain predicates ϕ and ψ , there is no difficulty. That the class of ψ 's includes the class of ϕ 's means simply that everything which is a ϕ is a ψ , which, as we saw in Chapter I, is a truth-function. But we have seen that mathematics has (at least apparently) to deal also with classes which are not given

by defining predicates. (Such classes occur not merely when mentioned separately, but also in any statement about "all classes", "all real numbers".) Let us take two such classes as simple as possible, the class (a, b, c) and the class (a, b) . Then that the class (a, b, c) includes the class (a, b) is, in a broad sense, tautological and apart from its triviality would be a mathematical proposition; but it does not seem to be a tautology in Wittgenstein's sense, that is a certain sort of truth-function of elementary propositions. The obvious way of trying to make it a truth-function is to introduce identity and write " (a, b) is contained in (a, b, c) " as " $(x) : .x = a . \vee .x = b : \supset : x = a . \vee .x = b . \vee .x = c$ ". This certainly looks like a tautological truth-function, whose ultimate arguments are values of " $x = a$ ", " $x = b$ ", " $x = c$ ", that is propositions like " $a = a$ ", " $b = a$ ", " $d = a$ ". But these are not real propositions at all; in " $a = b$ " either " a ", " b " are names of the same thing, in which case the proposition says nothing, or of different things, in which case it is absurd. In neither case is it the assertion of a fact; it only appears to be a real assertion by confusion with the case when " a " or " b " is not a name, but a description*. When " a ", " b " are both names, the only significance which can be placed on " $a = b$ " is that it indicates that we use " a ", " b " as names of the same thing or, more generally, as equivalent symbols.

The preceding and other considerations led Wittgenstein to the view that mathematics does not consist of tautologies, but of what he called "equations", for which I should prefer to substitute "identities". That is, formulae of the form " $a = b$ " where " a ", " b " are equivalent symbols. There is a certain plausibility in such an account of, for instance, " $2+2 = 4$ ". Since "I have $2+2$ hats", "I have 4 hats" are the same proposition†, " $2+2$ " and " 4 " are equivalent symbols. As it stands this is obviously a ridiculously narrow view of mathematics, which confines it to simple arithmetic, but it is interesting to see whether a theory of mathematics could not be constructed with identities for its foundation. I have spent a lot of time developing such a theory, and found that it was faced with what seemed to me insuperable difficulties. It would be out of place here to give a detailed survey of this blind alley, but I shall try to indicate in a general way the obstructions which block its end.

First of all we have to consider of what kind mathematical pro-

* For a fuller discussion of identity see the next chapter.

† In the sense explained in Chapter I. They clearly are not the same sentence, but they are the same truth-function of atomic propositions and so assert the same fact.

positions will on such a theory be. We suppose the most primitive type to be the identity " $a = b$ ", which only becomes a real proposition if it is taken to be about not the things meant by " a ", " b ", but these symbols themselves; mathematics then consists of propositions built up out of identities by a process analogous to that by which ordinary propositions are constructed out of atomic ones; that is to say, mathematical propositions are (on this theory), in some sense, truth-functions of identities. Perhaps this is an overstatement, and the theory might not assert all mathematical propositions to be of this form, but it is clearly one of the important forms that would be supposed to occur. Thus

$$"x^2 - 3x + 2 = 0 : \supset_x : x = 2 . \vee . x = 1"$$

would be said to be of this form, and would correspond to a verbal proposition which was a truth-function of the verbal propositions corresponding to the arguments " $x = 2$ ", etc. Thus the above proposition would amount to "If ' $x^2 - 3x + 2$ ' means 0, ' x ' means 2 or 1". Mathematics would then be, in part at least, the activity of constructing formulae which corresponded in this way to verbal propositions. Such a theory would be difficult and perhaps impossible to develop in detail, but there are, I think, other and simpler reasons for dismissing it. These arise as soon as we cease to treat mathematics as an isolated structure, and consider the mathematical elements in non-mathematical propositions. For simplicity let us confine ourselves to cardinal numbers, and suppose ourselves to know the analysis of the proposition that the class of ϕ 's is n in number ($\hat{x}(\phi x) \epsilon n$). Here ϕ may be any ordinary predicate defining a class, e.g. the class of ϕ 's may be the class of Englishmen. Now take such a proposition as "The square of the number of ϕ 's is greater by two than the cube of the number of ψ 's". This proposition we cannot, I think, help analysing in this sort of way.

$$(\exists m, n) . \hat{x}(\phi x) \epsilon m . \hat{x}(\psi x) \epsilon n . m^2 = n^3 + 2.$$

It is an empirical not a mathematical proposition, and is about the ϕ 's and ψ 's, not about symbols; yet there occurs in it the mathematical pseudo-proposition $m^2 = n^3 + 2$, of which, according to the theory under discussion, we can only make sense by taking it to be about symbols, thereby making the whole proposition partly about symbols. Moreover, being an empirical proposition, it is a truth-function of elementary propositions expressing agreement with those possibilities which give numbers of ϕ 's and ψ 's satisfying $m^2 = n^3 + 2$. Thus " $m^2 = n^3 + 2$ " is not, as it seems to be, one of the truth-arguments in the proposition

above, but rather part of the truth-function like “ \sim ” or “ \vee ” or “ \exists , m , n ”, which determine which truth-function of elementary propositions it is that we are asserting. Such a use of $m^2 = n^3 + 2$ the identity theory of mathematics is quite inadequate to explain.

On the other hand, the tautology theory would do everything which is required; according to it $m^2 = n^3 + 2$ would be a tautology for the values of m and n which satisfy it, and a contradiction for all others. So

$$\hat{x}(\phi x) \epsilon m . \hat{x}(\psi x) \epsilon n . m^2 = n^3 + 2$$

would for the first set of values of m , n be equivalent to

$$\hat{x}(\phi x) \epsilon m . \hat{x}(\psi x) \epsilon n$$

simply, “ $m^2 = n^3 + 2$ ” being tautologous, and therefore superfluous; and for all other values it would be self-contradictory. So that

$$“(\exists m, n) : \hat{x}(\phi x) \epsilon m . \hat{x}(\psi x) \epsilon n . m^2 = n^3 + 2”$$

would be the logical sum of the propositions

$$“\hat{x}(\phi x) \epsilon m . \hat{x}(\psi x) \epsilon n”$$

for all m, n satisfying $m^2 = n^3 + 2$, and of contradictions for all other m, n , and is therefore the proposition we require, since in a logical sum the contradictions are superfluous. So this difficulty, which seems fatal to the identity theory, is escaped altogether by the tautology theory, which we are therefore encouraged to pursue, and to see if we cannot find a way of overcoming the difficulties which we found would confront us in attempting to reduce an extensional calculus to a calculus of truth-functions. Such a solution is attempted in *Principia Mathematica*, and will be discussed in the next chapter; but before we proceed to this we must say something about the well known contradictions of the theory of aggregates which our theory will also have to escape.

It is not sufficiently remarked, and the fact is entirely neglected in *Principia Mathematica*, that these contradictions fall into two fundamentally distinct groups, which we will call A and B. The best known ones are divided as follows:—

- A. (1) The class of all classes which are not members of themselves.
 (2) The relation between two relations when one does not have itself to the other.
 (3) Burali Forti’s contradiction of the greatest ordinal.

B. (4) "I am lying".

(5) The least integer not nameable in fewer than nineteen syllables.

(6) The least undefinable ordinal.

(7) Richard's Contradiction.

(8) Weyl's Contradiction about "heterologisch"*.

The principle according to which I have divided them is of fundamental importance. Group A consists of contradictions, which, were no provision made against them, would occur in a logical or mathematical system itself. They involve only logical or mathematical terms such as class and number, and show that there must be something wrong with our logic or mathematics. But the contradictions of Group B are not purely logical, and cannot be stated in logical terms alone, for they all contain some reference to thought, language, or symbolism, which are not formal but empirical terms. So they may be due not to faulty logic or mathematics, but to faulty ideas concerning thought and language. If so, they would not be relevant to mathematics or to logic, if by "logic" we mean a symbolic system, though, of course, they would be relevant to logic in the sense of the analysis of thought†.

This view of the second group of contradictions is not original. For instance, Peano decided that "Exemplo de Richard non pertine ad Mathematica, sed ad linguistica"‡, and therefore dismissed it. But such an attitude is not completely satisfactory. We have contradictions involving both mathematical and linguistic ideas; the mathematician dismisses them by saying that the fault must lie in the linguistic elements, but the linguist may equally well dismiss it for the opposite reason, and the contradictions will never be solved. The only solution which has ever been given§, that in *Principia Mathematica*, definitely attributed the contradictions to bad logic, and it is up to opponents of this view to show clearly the fault in what Peano called linguistics, but

* For the first seven of these see *Principia Mathematica*, 1 (1910), 63. For the eighth see Weyl, *Das Kontinuum*, 2.

† These two meanings of "logic" are frequently confused. It really should be clear that those who say mathematics is logic are not meaning by "logic" at all the same thing as those who define logic as the analysis and criticism of thought.

‡ *Rivista di Mat.*, 8 (1906), 157.

§ Other so-called solutions are merely inadequate excuses for not giving a solution.

what I should prefer to call epistemology, to which these contradictions are due.

II. PRINCIPIA MATHEMATICA.

In the last chapter I tried to explain the difficulties which faced the theory that the propositions of mathematics are tautologies; in this we have to discuss the attempted solution of these difficulties given in *Principia Mathematica*. I shall try to show that this solution has three important defects, and the remainder of this essay will be devoted to expounding a modified theory from which these defects have been removed.

The theory of *Principia Mathematica* is that every class or aggregate (I use the words as synonyms) is defined by a propositional function—that is, consists of the values of x , for which " ϕx " is true, where " ϕx " is a symbol which expresses a proposition if any symbol of appropriate type be substituted for " x ". This amounts to saying that every class has a defining property. Let us take the class consisting of a and b ; why, it may be asked, must there be a function $\phi \hat{x}$ such that " ϕa ", " ϕb " are true, but all other " ϕx "'s false? This is answered by giving as such a function " $x = a \vee x = b$ ". Let us, for the present, neglect the difficulties connected with identity, and accept this answer; it shows us that any finite class is defined by a propositional function constructed by means of identity, but as regards infinite classes it leaves us exactly where we were before, that is, without any reason to suppose that they are all defined by propositional functions, for it is impossible to write down an infinite series of identities. To this it will be answered that a class can only be given to us either by enumeration of its members, in which case it must be finite, or by giving a propositional function which defines it. So that we cannot be in any way concerned with infinite classes or aggregates, if such there be, which are not defined by propositional functions.* But this argument contains a common mistake, for it supposes that, because we cannot consider a thing individually, we can have no concern with it at all. Thus, although an infinite indefinable class cannot be mentioned by itself, it is, nevertheless, involved in any statement beginning "All classes" or "There is a class such that", and if indefinable classes are excluded the meaning of all such statements will be fundamentally altered.

Whether there are indefinable classes or not is an empirical ques-

* For short, I shall call such classes "indefinable classes".

tion; both possibilities are perfectly conceivable. But even if, in fact, all classes are definable, we cannot in our logic identify classes with definable classes without destroying the apriority and necessity which is the essence of logic. But in case any one still thinks that by classes we mean definable classes, and by "there is a class", "there is a definable class," let him consider the following illustration. This illustration does not concern exactly this problem, but the corresponding problem for two variables, the existence of relations in extension not definable by propositional functions of two variables. But this question is clearly so analogous to the other that the answers to both must be the same.

Consider the proposition " $\hat{x}(\phi x) \text{ sm } \hat{x}(\psi x)$ " (*i.e.* the class defined by $\phi\hat{x}$ has the same cardinal as that defined by $\psi\hat{x}$); this is defined to mean that there is a one-one relation in extension whose domain is $\hat{x}(\phi x)$ and whose converse domain is $\hat{x}(\psi x)$. Now if by relation in extension we mean definable relation in extension, this means that two classes have the same cardinal only when there is a real relation or function $f(x, y)$ correlating them term by term. Whereas clearly what was meant by Cantor, who first gave this definition, was merely that the two classes were such that they could be correlated, not that there must be a propositional function which actually correlated them* Thus the classes of male and female angels may be infinite and equal in number, so that it would be possible to pair off completely the male with the female, without there being any real relation such as marriage correlating them. The possibility of indefinable classes and relations in extension is an essential part of the extensional attitude of modern mathematics which we emphasized in Chapter I, and that it is neglected in *Principia Mathematica* is the first of the three great defects in that work. The mistake is made not by having a primitive proposition, asserting that all classes are definable, but by giving a definition of class which applies only to definable classes, so that all mathematical propositions about some or all classes are misinterpreted. This misinterpretation is not merely objectionable on its own account in a general way, but is especially pernicious in connexion with the multiplicative axiom, which is a tautology when properly interpreted; but when misinterpreted after the fashion of *Principia Mathematica* becomes a significant empirical proposition, which there is no reason to suppose true. This will be shown in Chapter V.

The second defect in *Principia Mathematica* represents a failure to

* Cf. Johnson, *Logic*, 2 (1922), 159.

overcome, not, like the first, the difficulties raised by the extensionality of mathematics, but those raised by the contradictions discussed at the end of Chapter I. These contradictions it was proposed to remove by what is called the Theory of Types, which consists really of two distinct parts directed respectively against the two groups of contradictions. These two parts were unified by being both deduced in a rather sloppy way from the "vicious-circle principle", but it seems to me essential to consider them separately.

The contradictions of Group A are removed by pointing out that a propositional function cannot significantly take itself as argument, and by dividing functions and classes into a hierarchy of types according to their possible arguments. Thus the assertion that a class is a member of itself is neither true nor false, but meaningless. This part of the theory of types seems to me unquestionably correct, and I shall not discuss it further.

The first part of the theory, then, distinguishes types of propositional functions by their arguments; thus there are functions of individuals, functions of functions of individuals, functions of functions of functions of individuals, and so on. The second part designed to meet the second group of contradictions requires further distinctions between the different functions which take the same arguments, for instance between the different functions of individuals. The following explanation of these distinctions is based on the introduction to the second edition of *Principia Mathematica*.

We start with atomic propositions, which have been explained in Chapter I. Out of these by means of the stroke ($p/q =$ not both p and q are true) we can construct any truth-function of a finite number of atomic propositions as arguments. The assemblage of propositions so obtained are called elementary propositions. By substituting a variable for the name of an individual in one or more of its occurrences in an elementary proposition we obtain an elementary function of individuals. An *elementary function* of individuals, " $\hat{\phi}x$ " is therefore one whose values are elementary propositions, that is, truth-functions of a finite number of atomic propositions. Such functions were called, in the first edition of *Principia Mathematica*, *predicative functions*. We shall speak of them by their new name, and in the next chapter use "predicative function" in a new and original sense, for which it seems more appropriate. In general, an elementary function or "matrix" of one or more variables, whether these are individuals or not, is one whose values are elementary propositions. Matrices are denoted by a mark of exclamation after the functional symbol. Thus " $F!(\hat{\phi}! \hat{z}, \hat{\psi}! \hat{z}, \hat{x}, \hat{y})$ " is a matrix

having two individuals and two elementary functions of individuals as arguments.

From an elementary function " $\phi! \hat{x}$ " we obtain, as in Chapter I, the propositions " $(x). \phi! x$ " and " $(\exists x). \phi! x$ " which respectively assert the truth of all, and at least one of the values of " $\phi! x$ ". Similarly from an elementary function of two individuals $\phi!(\hat{x}, \hat{y})$ we obtain functions of one individual such as $(y). \phi!(\hat{x}, y)$, $(\exists y). \phi!(\hat{x}, y)$. The values of these functions are propositions such as $(y). \phi!(a, y)$ which are not elementary propositions; hence the functions themselves are not elementary functions. Such functions, whose values result from generalizing a matrix all whose values are individuals, are called first-order functions, and written $\phi_1 \hat{x}$.

Suppose a is a constant. Then " $\phi! a$ " will denote for the various values of ϕ , all the various elementary propositions of which a is a constituent. We can thus form the propositions $(\phi). \phi! a$, $(\exists \phi). \phi! a$ respectively asserting the truth of all, and at least one of the above assemblage of propositions. More generally we can assert by writing $(\phi). F!(\phi! \hat{z})$, $(\exists \phi). F!(\phi! \hat{z})$ the truth of all and at least one of the values of $F!(\phi! \hat{z})$. Such propositions are clearly not elementary, so that such a function as $(\phi). F!(\phi! \hat{z}, x)$ is not an elementary function of x . Such a function involving the totality of elementary functions is said to be of the second order and written $\phi_2 x$. By adopting the new variable ϕ_2 "we shall obtain other new functions

$$(\phi_2). f!(\phi_2 \hat{z}, x), \quad (\exists \phi_2). f!(\phi_2 \hat{z}, x),$$

which are again not among values for $\hat{\phi}_2 x$ (where ϕ_2 is the argument), because the totality of values of ϕ_2 , which is now involved, is different from the totality of values of $\phi!$, which was formerly involved. However much we may enlarge the meaning of ϕ , a function of x in which ϕ occurs as apparent variable has a correspondingly enlarged meaning so that, however ϕ may be defined, $(\phi). f!(\phi \hat{z}, x)$ and $(\exists \phi). f!(\phi \hat{z}, x)$ can never be values for $\hat{\phi} x$. To attempt to make them so is like attempting to catch one's own shadow. It is impossible to obtain one variable which embraces among its values all possible functions of individuals"*.

For the way in which this distinction of functions into orders of which no totality is possible is used to escape the contradictions of Group B, which are shown to result from the ambiguities of language which disregard this distinction, reference may be made to *Principia*

* *Principia Mathematica*, 1 (1925), 2 ed., xxxiv.

*Mathematica**. Here it may be sufficient to apply the method to a contradiction not given in that work, which is peculiarly free from irrelevant elements; I mean, Weyl's contradiction concerning "heterologisch"†, which must now be explained. Some adjectives have meanings which are predicates of the adjective word itself; thus the word "short" is short, but the word "long" is not long. Let us call adjectives whose meanings are predicates of them, like "short", autological; others heterological. Now is "heterological" heterological? If it is, its meaning is not a predicate of it, that is, it is not heterological. But if it is not heterological, its meaning is a predicate of it, and therefore it is heterological. So we have a complete contradiction.

According to the principles of *Principia Mathematica*, this contradiction would be solved in the following way. An adjective word is the symbol for a propositional function, e.g. " ϕ " for $\phi\hat{x}$. Let R be the relation of meaning between " ϕ " and $\phi\hat{x}$. Then " w is heterological" is " $(\exists\phi). wR(\phi\hat{x}). \sim\phi w$ ". In this, as we have seen, the apparent variable ϕ must have a definite range of values (e.g. the range of elementary functions), of which $Fx = \therefore (\exists\phi): xR(\phi\hat{x}). \sim\phi x$ cannot itself be a member. So that "heterological" or " F " is not itself an adjective in the sense in which " ϕ " is. We do not have $(\exists\phi). "F" R(\phi\hat{x})$ because the meaning of " F " is not a function included in the range of " ϕ ". So that when heterological and autological are unambiguously defined, "heterological" is not an adjective in the sense in question, and is neither heterological nor autological, and there is no contradiction.

Thus this theory of a hierarchy of orders of functions of individuals escapes the contradictions, but it lands us in an almost equally serious difficulty, for it invalidates many important mathematical arguments which appear to contain exactly the same fallacy as the contradictions. In the first edition of *Principia Mathematica* it was proposed to justify these arguments by a special axiom, the axiom of reducibility, which asserted that to every non-elementary function there is an equivalent elementary function‡. This axiom there is no reason to suppose true, and if it were true, this would be a happy accident and not a logical necessity, for it is not a tautology. This will be shown positively in Chapter V; but for the present it should be sufficient that it does not seem to be a tautology and that there is no reason to suppose that it is one. Such an axiom

* *Principia Mathematica*, 1 ed., 1 (1910), 117.

† Weyl, *Das Kontinuum*, 2.

‡ Two functions are called equivalent when the same arguments render them both true or both false. (German *umfangsgleich*.)

has no place in mathematics, and anything which cannot be proved without using it cannot be regarded as proved at all.

It is perhaps worth while, parenthetically, to notice a point which is sometimes missed. Why, it may be asked, does not the axiom of reducibility reproduce the contradictions, which the distinction between elementary and other functions avoided? For it asserts that to any non-elementary there is an equivalent elementary function, and so may appear to lose again whatever was gained by making the distinction. This is not, however, the case, owing to the peculiar nature of the contradictions in question; for, as pointed out above, this second set of contradictions are not purely mathematical, but all involve the ideas of thought or meaning, in connexion with which equivalent functions (in the sense of equivalent explained above) are not interchangeable; for instance, one can be meant by a certain word or symbol, but not the other; and one can be definable, and not the other*. On the other hand, any purely mathematical contradiction which arose from confusing elementary and non-elementary functions would be reinstated by the axiom of reducibility, owing to the extensional nature of mathematics, in which equivalent functions are interchangeable. But no such contradiction has been shown to arise, so that the axiom of reducibility does not seem to be self-contradictory. These considerations bring out clearly the peculiarity of this second group of contradictions, and make it even more probable that they have a psychological or epistemological and not a purely logical or mathematical solution; so that there is something wrong with the account of the matter given in *Principia*.

The principal mathematical methods which appear to require the axiom of reducibility are mathematical induction and Dedekindian section, the essential foundations of arithmetic and analysis respectively. Mr. Russell has succeeded in dispensing with the axiom in the first case †, but holds out no hope of a similar success in the second. Dedekindian section is thus left as an essentially unsound method, as has often been

* Dr. L. Chwistek appears to have overlooked this point that, if a function is definable, the equivalent elementary function need not also be definable in terms of given symbols. In his paper "Über die Antinomien der Prinzipien der Mathematik" in *Math. Zeitschrift*, 14 (1922), 236-243, he denotes by S a many-one relation between the natural numbers and the classes defined by functions definable in terms of certain symbols. $\phi\hat{z}$ being a non-elementary function of this kind, he concludes that there must be an n , such that $nS\hat{z}(\phi z)$. This is, however, a fallacy since $nS\hat{z}(\phi z)$ means by definition

$$(\exists \psi) : \psi ! x \equiv_r \phi x . nS(\psi ! \hat{z})$$

and since $\psi ! \hat{z}$ is not necessarily definable in terms of the given symbols there is no reason for there being any such n .

† See *Principia Mathematica*, 1 (1925), 2 ed., Appendix B.

emphasized by Weyl*, and ordinary analysis crumbles into dust. That these are its consequences is the second defect in the theory of *Principia Mathematica*, and, to my mind, an absolutely conclusive proof that there is something wrong. For as I can neither accept the axiom of reducibility nor reject ordinary analysis, I cannot believe in a theory which presents me with no third possibility.

The third serious defect in *Principia Mathematica* is the treatment of identity. It should be explained that what is meant is numerical identity, identity in the sense of counting as one, not as two. Of this the following definition is given :

$$"x = y. = : (\phi) : \phi! x. \supset . \phi! y : \text{Df} " \dagger$$

that is, two things are identical if they have all their elementary properties in common.

In *Principia* this definition is asserted to depend on the axiom of reducibility, because, apart from this axiom, two things might have all their elementary properties in common, but still disagree in respect of functions of higher order, in which case they could not be regarded as numerically identical‡. Although, as we shall see, the definition is to be rejected on other grounds, I do not think it depends in this way on the axiom of reducibility. For though rejecting the axiom of reducibility destroys the obvious general proof that two things agreeing in respect of all elementary functions agree also in respect of all other functions, I think that this would still follow and could probably be proved in any particular case. For example, take a typical function of the second order

$$(\phi) : f!(\phi! \hat{z}, x), \quad (\exists \phi) . f!(\phi! \hat{z}, x).$$

$$\text{Then, if we have} \quad (\phi) : \phi! x. \equiv . \phi! y \quad (x = y),$$

$$\text{it follows that} \quad (\phi) : f!(\phi! \hat{z}, x). \equiv . f!(\phi! \hat{z}, y),$$

because $f!(\phi! \hat{z}, x)$ is an elementary function of x . Whence

$$(\phi) . f!(\phi! \hat{z}, x) : \equiv : (\phi) . f!(\phi! \hat{z}, y)$$

$$\text{and} \quad (\exists \phi) . f!(\phi! \hat{z}, x) : \equiv : (\exists \phi) . f!(\phi! \hat{z}, y).$$

* See H. Weyl, *Das Kontinuum*, and "Ueber die neue Grundlagenkrise der Mathematik", *Math. Zeitschrift*, 10 (1921), 39-79.

† 13. 01.

‡ *Principia Mathematica*, 1 ed., 1 (1910), 177.

Hence rejecting the axiom of reducibility does not immediately lead to rejecting the definition of identity.

The real objection to this definition of identity is the same as that urged above against defining classes to be definable classes, that it is a misinterpretation, that it does not define the meaning with which the symbol for identity is actually used. This can be easily seen in the following way: the definition makes it self-contradictory for two things to have all their elementary properties in common. Yet this is really perfectly possible, even if, in fact, it never happens. Take two things, a and b . Then there is nothing self-contradictory in a having any self-consistent set of elementary properties, nor in b having this set, nor therefore, obviously, in both a and b having them, nor therefore in a and b having all their elementary properties in common. Hence, since this is logically possible, it is essential to have a symbolism which allows us to consider this possibility and does not exclude it by definition.

It is futile to raise the objection that it is not possible to distinguish two things which have all their properties in common, since to give them different names would imply that they had the different properties of having those names. For although this is perfectly true, that is to say, I cannot, for the reason given, know of any two particular indistinguishable things, yet I can perfectly well consider the possibility, or even know that there are two indistinguishable things, without knowing which they are. To take an analogous situation: since there are more people on the earth than hairs on any one person's head, I know that there must be at least two people with the same number of hairs, but I do not know which two people they are.

These arguments are reinforced by Wittgenstein's discovery that the sign of identity is not a necessary constituent of logical notation, but can be replaced by the convention that different signs must have different meanings. This will be found in *Tractatus Logico-Philosophicus*, p. 139; the convention is slightly ambiguous, but it can be made definite, and is then workable, although generally inconvenient. But even if of no other value, it provides an effective proof that identity can be replaced by a symbolic convention, and is therefore no genuine propositional function, but merely a logical device.

We conclude, therefore, that the treatment of identity in *Principia Mathematica* is a misinterpretation of mathematics, and just as the mistaken definition of classes is particularly unfortunate in connexion with the Multiplicative Axiom, so the mistaken definition of identity is especially misleading with regard to the Axiom of Infinity; for the two propositions, "There are an infinite number of things" and "There are

an infinite number of things differing from one another with regard to elementary functions'', are, as we shall see in Chapter V, extremely different.

III. PREDICATIVE FUNCTIONS.

In this chapter we shall consider the second of the three objections which we made in the last chapter to the theory of the foundations of mathematics given in *Principia Mathematica*. This objection, which is perhaps the most serious of the three, was directed against the "Theory of Types," which seemed to involve either the acceptance of the illegitimate Axiom of Reducibility, or the rejection of such a fundamental type of mathematical argument as Dedekind section. We saw that this difficulty came from the second of the two parts into which the theory was divided, namely that part which concerned the different ranges of functions of given arguments, *e.g.* individuals, and we have to consider whether this part of the theory of types cannot be amended so as to get out of the difficulty. We shall see that this can be done in a simple and straightforward way, which is a natural consequence of the logical theories of Mr. Wittgenstein.

We shall start afresh from part of his theory of propositions, of which something was said in the first chapter. We saw there that he explains propositions in general by reference to atomic propositions, every proposition expressing agreement and disagreement with truth-possibilities of atomic propositions. We saw also that we could construct many different symbols all expressing agreement and disagreement with the same sets of possibilities. For instance,

$$"p \supset q", \quad "\sim p \vee q", \quad "\sim : p \sim q", \quad "\sim q \supset \sim p"$$

are such a set, all agreeing with the three possibilities

$$"p \cdot q", \quad "\sim p \cdot q", \quad "\sim p \cdot \sim q",$$

but disagreeing with " $p \sim q$ ". Two symbols of this kind, which express agreement and disagreement with the same sets of possibilities, are said to be instances of the same proposition. They are instances of it just as all the "the"'s on a page are instances of the word "the". But whereas the "the"'s are instances of the same word on account of their physical similarity, different symbols are instances of the same proposition because they have the same sense, that is, express agreement with the same sets of possibilities. When we speak of propositions

we shall generally mean the types of which individual symbols are instances, and we shall include types of which there may be no instances. This is inevitable, since it cannot be any concern of ours whether anyone has actually symbolized or asserted a proposition, and we have to consider all propositions in the sense of all possible assertions whether or not they have been asserted.

Any proposition expresses agreement and disagreement with complementary sets of truth-possibilities of atomic propositions; conversely given any set of these truth-possibilities, it would be logically possible to assert agreement with them and disagreement with all others, and the set of truth-possibilities therefore determines a proposition. This proposition may in practice be extremely difficult to express through the poverty of our language, for we lack both names for many objects and methods of making assertions involving an infinite number of atomic propositions, except in relatively simple cases, such as " $(x) . \phi x$ ", which involves the (probably) infinite set of (in certain cases) atomic propositions, " ϕa ", " ϕb ", &c. Nevertheless, we have to consider propositions which our language is inadequate to express. In " $(x) . \phi x$ " we assert the truth of all possible propositions which would be of the form " ϕx " whether or not we have names for all the values of x . General propositions must obviously be understood as applying to everything, not merely to everything for which we have names.

We come now to a most important point in connection with the Theory of Types. We explained in the last chapter what was meant by an elementary proposition, namely one constructed explicitly as a truth-function of atomic propositions. We have now to see that, on the theory of Wittgenstein, elementary is not an adjective of the proposition type at all, but only of its instances. For an elementary and a non-elementary propositional symbol could be instances of the same proposition. Thus suppose a list was made of all individuals as " a ", " b ", ..., " z ". Then, if $\hat{\phi x}$ were an elementary function, " $\phi a . \phi b \dots \phi z$ " would be an elementary proposition, but " $(x) . \phi x$ " non-elementary; but these would express agreement and disagreement with the same possibilities and therefore be the same proposition. Or to take an example which could really occur, " ϕa " and " $\phi a : (\exists x) . \phi x$ ", which are the same proposition, since $(\exists x) . \phi x$ adds nothing to ϕa . But the first is elementary, the second non-elementary.

Hence some instances of a proposition can be elementary, and others non-elementary; so that elementary is not really a characteristic of the proposition, but of its mode of expression. "Elementary proposition" is like "spoken word"; just as the same word can be both spoken and

written, so the same proposition can be both elementarily and non-elementarily expressed.

After these preliminary explanations we proceed to a theory of propositional functions. By a propositional function of individuals we mean a symbol of the form " $f(\hat{x}, \hat{y}, \hat{z}, \dots)$ " which is such that, were the names of any individuals substituted for " \hat{x} ", " \hat{y} ", " \hat{z} ", ... in it, the result would always be a proposition. This definition needs to be completed by the explanation that two such symbols are regarded as the same function when the substitution of the same set of names in the one and in the other always gives the same proposition. Thus if " $f(a, b, c)$ " " $g(a, b, c)$ " are the same proposition for any set of a, b, c , $f(\hat{x}, \hat{y}, \hat{z})$ and $g(\hat{x}, \hat{y}, \hat{z})$ are the same function, even if they are quite different to look at.

A function* " $\phi\hat{x}$ " gives us for each individual a proposition, in the sense of a proposition type (which may not have any instances, for we may not have given the individual a name). So the function collects together a set of propositions, whose logical sum and product we assert by writing respectively " $(\exists x) . \phi x$ ", " $(x) . \phi x$ ". This procedure can be extended to the case of several variables. Consider " $\phi(\hat{x}, \hat{y})$ "; give y any constant value η , and " $\phi(\hat{x}, \eta)$ " gives a proposition when any individual name is substituted for \hat{x} , and is therefore a function of one variable, from which we can form the propositions

$$"(\exists x) . \phi(x, \eta)", \quad "(x) . \phi(x, \eta)".$$

Consider next " $(\exists x) . \phi(x, \hat{y})$ "; this, as we have seen, gives a proposition when any name (*e.g.* " η ") is substituted for " \hat{y} ", and is therefore a function of one variable from which we can form the propositions

$$(\exists y) : (\exists x) . \phi(x, y) \quad \text{and} \quad (y) : (\exists x) . \phi(x, y).$$

As so far there has been no difficulty, we shall attempt to treat functions of functions in exactly the same way as we have treated functions of individuals. Let us take, for simplicity, a function of one variable which is a function of individuals. This would be a symbol of the form " $f(\hat{\phi}\hat{x})$ ", which becomes a proposition on the substitution for " $\hat{\phi}\hat{x}$ " of any function of an individual. " $f(\phi\hat{x})$ " then collects together

* By "function" we shall in future always mean propositional function unless the contrary is stated.

a set of propositions, one for each function of an individual, of which we assert the logical sum and product by writing respectively " $(\exists \phi) \cdot f(\hat{\phi}x)$ ", " $(\phi) \cdot f(\hat{\phi}x)$ ".

But this account suffers from an unfortunate vagueness as to the range of functions $\hat{\phi}x$ giving the values of $f(\hat{\phi}x)$ of which we assert the logical sum or product. In this respect there is an important difference between functions of functions and functions of individuals which is worth examining closely. It appears clearly in the fact that the expressions "function of functions", "function of individuals" are not strictly analogous; for, whereas functions are symbols, individuals are objects, so that to get an expression analogous to "function of functions" we should have to say "function of names of individuals". On the other hand, there does not seem any simple way of altering "function of functions" so as to make it analogous to "function of individuals", and it is just this which causes the trouble. For the range of values of a function of individuals is definitely fixed by the range of individuals, an objective totality which there is no getting away from. But the range of arguments to a function of functions is a range of symbols, all symbols which become propositions by inserting in them the name of an individual. And this range of symbols, actual or possible, is not objectively fixed, but depends on our methods of constructing them and requires more precise definition.

This definition can be given in two ways, which may be distinguished as the subjective and objective methods. The subjective* method is that adopted in *Principia Mathematica*; it consists in defining the range of functions as all those which could be constructed in a certain way, in the first instance by sole use of the "/" sign. We have seen how it leads to the impasse of the Axiom of Reducibility. I, on the other hand, shall adopt the entirely original objective method which will lead us to a satisfactory theory in which no such axiom is required. This method is to treat functions of functions as far as possible in the same way as functions of individuals. The signs which can be substituted as arguments in " $\hat{\phi}x$ ", a function of individuals, are determined by their meanings; they must be names of individuals. I propose similarly to determine the symbols which can be substituted as arguments in " $f(\hat{\phi}x)$ " not by the manner of their construction, but by their meanings. This is more difficult, because functions do not mean single objects as names do, but have meaning in a more complicated way derived from the meanings

* I do not wish to press this term; I merely use it because I can find no better.

of the propositions which are their values. The problem is ultimately to fix as values of $f(\phi\hat{x})$ some definite set of propositions so that we can assert their logical product and sum. In *Principia Mathematica* they are determined as all propositions which can be constructed in a certain way. My method, on the other hand, is to disregard how we could construct them, and determine them by a description of their senses or imports; and in so doing we may be able to include in the set propositions which we have no way of constructing, just as we include in the range of values of ϕx propositions which we cannot express from lack of names for the individuals concerned.

We must begin the description of the new method with the definition of an atomic function of individuals, as the result of replacing by variables any of the names of individuals in an atomic proposition expressed by using names alone; where if a name occur more than once in the proposition it may be replaced by the same or different variables, or left alone in its different occurrences. The values of an atomic function of individuals are thus atomic propositions.

We next extend to propositional functions the idea of a truth-function of propositions. (At first, of course, the functions to which we extend it are only atomic, but the extension works also in general, and so I shall state it in general.) Suppose we have functions $\phi_1(\hat{x}, \hat{y})$, $\phi_2(\hat{x}, \hat{y})$, etc., then by saying that a function $\psi(\hat{x}, \hat{y})$ is a certain truth-function (e.g. the logical sum) of the functions $\phi_1(\hat{x}, \hat{y})$, $\phi_2(\hat{x}, \hat{y})$, etc., and the propositions p , q , &c., we mean that any value of $\psi(x, y)$, say $\psi(a, b)$, is that truth-function of the corresponding values of $\phi_1(x, y)$, $\phi_2(x, y)$, etc., i.e. $\phi_1(a, b)$, $\phi_2(a, b)$, etc., and the propositions p , q , etc. This definition enables us to include functions among the arguments of any truth-function, for it always give us a unique function which is that truth-function of those arguments, e.g. the logical sum of $\phi_1(\hat{x})$, $\phi_2(\hat{x})$, ... is determined as $\psi(x)$, where $\psi(a)$ is the logical sum of $\phi_1 a$, $\phi_2 a$, ..., a definite proposition for each a , so that $\psi(x)$ is a definite function. It is unique because, if there were two, namely $\psi_1(x)$ and $\psi_2(x)$, $\psi_1(a)$ and $\psi_2(a)$ would for each a be the same proposition, and hence the two functions would be identical.

We can now give the most important definition in this theory, that of a predicative function. I do not use this term in the sense of *Principia Mathematica*, 1st ed., for which I follow Mr. Russell's later work in using "elementary". The notion of a predicative function, in my sense, is one which does not occur in *Principia*, and marks the essential divergence of the two methods of procedure. A predicative function of

individuals is one which is any truth-function of arguments which, whether finite or infinite in number, are all either atomic functions of individuals or propositions*. This defines a definite range of functions of individuals which is wider than any range occurring in *Principia*. It is essentially dependent on the notion of a truth-function of an infinite number of arguments; if there could only be a finite number of arguments our predicative functions would be simply the elementary functions of *Principia*. Admitting an infinite number involves that we do not define the range of functions as those which could be constructed in a certain way, but determine them by a description of their meanings. They are to be truth-functions, not explicitly in their appearance, but in their significance, of atomic functions and propositions. In this way we shall include many functions which we have no way of constructing, and many which we construct in quite different ways. Thus, supposing $\phi(\hat{x}, \hat{y})$ is an atomic function, p a proposition,

$$\phi(\hat{x}, \hat{y}), \quad \phi(\hat{x}, \hat{y}) \cdot \vee \cdot p, \quad (y) \cdot \phi(\hat{x}, y)$$

are all predicative functions. [The last is predicative because it is the logical product of the atomic functions $\phi(\hat{x}, y)$ for different values of y .]

For functions of functions there are more or less analogous definitions. First, an atomic function of (predicative†) functions of individuals and of individuals can only have one functional argument, say ϕ , but may have many individual arguments, $x, y, \&c.$, and must be of the form $\phi(x, y, \dots, a, b, \dots)$ where "a", "b", ... are names of individuals. In particular, an atomic function $f(\phi\hat{z})$ is of the form ϕa . A predicative function of (predicative) functions of individuals and of individuals is one which is a truth-function whose arguments are all either propositions or atomic functions of functions of individuals and of individuals

e.g. $\hat{\phi} a \cdot \supset \cdot \hat{\psi} b : \vee : p$ (a function of ϕ, ψ),

(x). $\hat{\phi} x$, the logical product of the atomic functions $\hat{\phi} a, \hat{\phi} b$, etc.

It is clear that a function only occurs in a predicative function through

* Before "propositions" we could insert "atomic" without narrowing the sense of the definition. For any proposition is a truth-function of atomic propositions, and a truth-function of a truth-function is again a truth-function.

† I put "predicative" in brackets because the definitions apply equally to the non-predicative functions dealt with in the next chapter.

its values. In this way we can proceed to define predicative functions of functions of functions and so on to any order.

Now consider such a proposition as $(\phi).f(\hat{\phi}\hat{x})$ where $f(\hat{\phi}\hat{x})$ is a predicative function of functions. We understand the range of values of ϕ to be all predicative functions; *i.e.* $(\phi).f(\hat{\phi}\hat{x})$ is the logical product of the propositions $f(\hat{\phi}\hat{x})$ for each predicative function, and as this is a definite set of propositions, we have attached to $(\phi).f(\hat{\phi}\hat{x})$ a definite significance.

Now consider the function of x , $(\phi).f(\hat{\phi}\hat{z}, x)$. Is this a predicative function? It is the logical product of the propositional functions of x , $f(\hat{\phi}\hat{z}, x)$ for the different ϕ 's which, since f is predicative, are truth-functions of ϕx and propositions possibly variable in ϕ but constant in x (*e.g.* ϕa). The ϕx 's, since the ϕ 's are predicative, are truth-functions of atomic functions of x . Hence the propositional functions of x , $f(\hat{\phi}\hat{z}, x)$ are truth-functions of atomic functions of x and propositions. Hence they are predicative functions, and therefore their logical product $(\phi).f(\hat{\phi}\hat{z}, x)$ is predicative. More generally it is clear that by generalization, whatever the type of the apparent variable, we can never create non-predicative functions; for the generalization is a truth-function of its instances, and, if these are predicative, so is it.

Thus all the functions of individuals which occur in *Principia* are in our sense predicative and included in our variable ϕ , so that all need for an axiom of reducibility disappears.

But, it will be objected, surely in this there is a vicious circle; you cannot include $F\hat{x} = (\phi).f(\hat{\phi}\hat{z}, \hat{x})$ among the ϕ 's, for it presupposes the totality of the ϕ 's. This is not, however, really a vicious circle. The proposition Fa is certainly the logical product of the propositions $f(\hat{\phi}\hat{z}, a)$, but to express it like this (which is the only way we can) is merely to describe it in a certain way, by reference to a totality of which it may be itself a member, just as we may refer to a man as the tallest in a group, thus identifying him by means of a totality of which he is himself a member without there being any vicious circle. The proposition Fa in its significance, that is the fact it asserts to be the case does not involve the totality of functions; it is merely our symbol which involves it. To take a particularly simple case, $(\phi).\phi a$ is the logical product of the propositions ϕa , of which it is itself one; but this is no more remarkable and no more vicious than is the fact that $p . q$ is the logical product of the set p, q , $p . q$, of which it is itself a member. The only difference is that, owing to our inability to write propositions

of infinite length, which is logically a mere accident, $(\phi) . \phi a$ cannot, like $p . q$, be elementarily expressed, but must be expressed as the logical product of a set of which it is also a member. If we had infinite resources and could express all atomic functions as $\psi_1 x, \psi_2 x$, then we could form all the propositions ϕa , that is all the truth-functions of $\psi_1 a, \psi_2 a$, etc., and among them would be one which was the logical product of them all, including itself, just as $p . q$ is the product of $p, q, p \vee q, p . q$. This proposition, which we cannot express directly, that is elementarily, we express indirectly as the logical product of them all by writing " $(\phi) . \phi a$ ". This is certainly a circuitous process, but there is clearly nothing vicious about it.

In this lies the great advantage of my method over that of *Principia Mathematica*. In *Principia* the range of ϕ is that of functions which can be elementarily expressed, and since $(\phi) . f(\phi ! \hat{x}, x)$ cannot be so expressed it cannot be a value of $\phi !$; but I define the values of ϕ not by how they can be expressed, but by what sort of senses their values have, or, rather, by how the facts their values assert are related to their arguments. I thus include functions which could not even be expressed by us at all, let alone elementarily, but only by a being with an infinite symbolic system. And any function formed by generalization being actually predicative, there is no longer any need for an axiom of reducibility.

It remains to show that my notion of predicative functions does not involve us in any contradictions. The relevant contradictions, as I have remarked before, all contain some word like "means", and I shall show that they are due to an essential ambiguity of such words and not to any weakness in the notion of a predicative function.

Let us take first Weyl's contradiction about "heterological", which we discussed in the last chapter. It is clear that the solution given there is no longer available to us. For, as before, if R is the relation of meaning between " ϕ " and $\phi \hat{x}$, " x is heterological" is equivalent to " $(\exists \phi) . xR(\phi \hat{z}) . \sim \phi x$ ", the range of ϕ being here understood to be that of predicative functions. Then

$$(\exists \phi) : xR(\phi \hat{z}) . \sim \phi x,$$

which I will call Fx , is itself a predicative function.

So

$$"F" R(F \hat{x})$$

and

$$(\exists \phi) : "F" R(\phi \hat{x}),$$

and therefore

$$F("F") \equiv \sim F("F"),$$

which is a contradiction.

It will be seen that the contradiction essentially depends on deducing $(\exists \phi): "F" R(\phi \hat{x})$ from $"F" R(F \hat{x})$. According to *Principia Mathematica* this deduction is illegitimate because $F \hat{x}$ is not a possible value of $\phi \hat{x}$. But if the range of $\phi \hat{x}$ is that of predicative functions, this solution fails, since $F \hat{x}$ is certainly a predicative function. But there is obviously another possible solution—to deny $"F" R(F \hat{x})$ the premiss of the deduction. $"F" R(F \hat{x})$ says that $"F"$ means $F \hat{x}$. Now this is certainly true for some meaning of "means", so to uphold our denial of it we must show some ambiguity in the meaning of meaning, and say that the sense in which $"F"$ means $F \hat{x}$, *i.e.* in which "heterological" means heterological, is not the sense denoted by $"R"$, *i.e.* the sense which occurs in the definition of heterological. We can easily show that this is really the case, so that the contradiction is simply due to an ambiguity in the word "meaning" and has no relevance to mathematics whatever.

First of all to speak of $"F"$ as meaning $F \hat{x}$ at all must appear very odd in view of our definition of a propositional function as itself a symbol. But the expression is merely elliptical. The fact which we try to describe in those terms is that we have arbitrarily chosen the letter $"F"$ for a certain purpose, so that $"Fx"$ shall have a certain meaning (depending on x). As a result of this choice $"F"$, previously non-significant, becomes significant; it has meaning. But it is clearly an impossible simplification to suppose that there is a single object F , which it means. Its meaning is more complicated than that, and must be further investigated.

Let us take the simplest case, an atomic proposition fully written out $"aSb"$, where $"a"$, $"b"$ are names of individuals and $"S"$ the name of a relation. Then $"a"$, $"b"$, $"S"$ mean in the simplest way the separate objects a , b , and S . Now suppose we define

$$\phi x = .aSx \quad \text{Df.}$$

Then $"\phi"$ is substituted for $"aS"$ and does not mean a single object, but has meaning in a more complicated way, in virtue of a three-termed relation to both a and S . Then we can say $"\phi"$ means $aS \hat{x}$, meaning by this that $"\phi"$ has this relation to a and S . We can extend this

account to deal with any elementary function, that is, to say that " $\phi!$ " means $\phi!\hat{x}$ means that " $\phi!$ " is related in a certain way to the objects $a, b, \text{etc.}$, involved in " $\phi!\hat{x}$ ".

But suppose now we take a non-elementary functional symbol, for example,

$$\phi_1 x := (y). yRx \quad \text{Df.}$$

Here the objects involved in $\phi_1\hat{x}$ include all individuals as values of y . And it is clear that " ϕ_1 " is not related to them in at all the same way, as " $\phi!$ " is to the objects in its meaning. For " $\phi!$ " is related to $a, b, \text{etc.}$, by being short for an expression containing names of $a, b, \text{etc.}$. But " ϕ_1 " is short for an expression not containing " a ", " b ", ..., but only an apparent variable, of which these can be values. Clearly " ϕ_1 " means what it means in quite a different and more complicated way from that in which " $\phi!$ " means. Of course, just as elementary is not really a characteristic of the proposition, it is not really a characteristic of the function; that is to say $\phi_1\hat{x}$ and $\phi!\hat{x}$ may be the same function, because $\phi_1 x$ is always the same proposition as $\phi!x$. Then " ϕ_1 ", " $\phi!$ " will have the same meaning, but will mean it, as we saw above, in quite different senses of meaning. Similarly a " ϕ_2 " which involves functional apparent variable will mean in a different and more complicated way still*.

Hence in the contradiction which we were discussing, if " R ", the symbol of the relation of meaning between " ϕ " and $\phi\hat{x}$, is to have any definite meaning, " ϕ " can only be a symbol of a certain type meaning in a certain way; suppose we limit " ϕ " to be an elementary function, by taking R to be the relation between " $\phi!$ " and $\phi!\hat{x}$.

Then " Fx " or " $(\exists\phi): xR(\phi z). \sim \phi x$ " is not elementary, but is a " ϕ_2 ".

Hence " F " means not in the sense of meaning denoted by " R " appropriate to " $\phi!$ "s, but in that appropriate to a " ϕ_2 ", so that we have $\sim : "F" R(F\hat{x})$, which, as we explained above, solves the contradiction for this case.

The essential point to understand is that the reason why

$$(\exists\phi) : "F" R(\phi\hat{x})$$

* Here the range of the apparent variable in " ϕ_2 " is the set of predicative functions not as in *Principia Mathematica* the set of elementary functions.

can only be true if " F " is an elementary function, is not that the range of ϕ is that of elementary functions, but that a symbol cannot have R to a function unless it (the symbol) is elementary. The limitation comes not from " $\exists\phi$ ", but from " R ". The distinctions of " $\phi!$ "s, " ϕ_1 "s, and " ϕ_2 "s apply to the symbols and to how they mean but not to what they mean. Therefore I always (in this section) enclosed " $\phi!$ ", " ϕ_1 " and " ϕ_2 " in inverted commas.

But it may be objected that this is an incomplete solution; for suppose we take for R the sum of the relations appropriate to " $\phi!$ "s, " ϕ_1 "s, and " ϕ_2 "s. Then " F ", since it still only contains $\exists\phi^*$, is still a " ϕ_2 ", and we must have in this case " F " $R(F\hat{x})$; which destroys our solution.

But this is not so because the extra complexity involved in the new R makes " F " not a " ϕ_2 ", but a more complicated symbol still. For with this new R , for which " ϕ_2 " $R(\phi_2\hat{x})$, since " ϕ_2x " is of some such form as $(\exists\phi).f(\phi\hat{z}, x)$, in $(\exists\phi). "F" R(\phi\hat{x})$ is involved at least a variable function $f(\phi\hat{z}, x)$ of functions of individuals, for this is involved in the notion of a variable " ϕ_2 ", which is involved in the variable ϕ taken in conjunction with R . For if anything has R to the predicative function $\phi\hat{x}$, $\phi\hat{x}$ must be expressible by either a " $\phi!$ " or a " ϕ_1 " or a " ϕ_2 ".

Hence $(\exists\phi). "F" R(\phi\hat{x})$ involves not merely the variable ϕ (predicative function of an individual) but also a hidden variable f (function of a function of an individual and an individual). Hence " Fx " or $(\exists\phi). xR(\phi\hat{x}). \sim\phi x$ is not a " ϕ_2 ", but what we may call a " ϕ_3 ", i.e. a function of individuals involving a variable function of functions of individuals. (This is, of course, not the same thing as a " ϕ_3 " in the sense of *Principia Mathematica*, 2nd edition.) Hence " F " means in a more complicated way still not included in R and we do not have " F " $R(F\hat{x})$ so that the contradiction again disappears.

What appears clearly from the contradictions is that we cannot obtain an all-inclusive relation of meaning for propositional functions. Whatever one we take there is still a way of constructing a symbol to mean in a way not included in our relation. The meanings of meaning form an illegitimate totality.

By the process begun above we obtain a hierarchy of propositions and a hierarchy of functions of individuals. Both are based on the fundamental hierarchy of individuals, functions of individuals, functions of functions of individuals, etc. A function of individuals we will call

* The range of ϕ in $\exists\phi$ is that of predicative functions including all " ϕ_1 "s, " ϕ_2 "s, etc., so it is not altered by changing R .

a function of type 1; a function of functions of individuals, a function of type 2, and so on.

We now construct the hierarchy of propositions as follows:—

Propositions of order 0 (elementary), containing no apparent variable.

„ „ „ 1, containing an individual apparent variable.

„ „ „ 2, containing an apparent variable whose values are functions of type 1.

„ „ „ n , containing an apparent variable whose values are functions of type $n-1$.

From this hierarchy we deduce another hierarchy of functions, irrespective of their types, according to the order of their values.

Thus functions of order 0 (matrices) contain no apparent variable;

„ „ „ „ 1 contain an individual apparent variable;

and so on; *i.e.* the values of a function of order n being propositions of order n . For this classification the types of the functions are immaterial.

We must emphasize the essential distinction between order and type. The type of a function is a real characteristic of it depending on the arguments it can take; but the order of a proposition or function is not a real characteristic, but what Peano called a pseudo-function. The order of a proposition is like the numerator of a fraction. Just as from " $x = y$ " we cannot deduce that the numerator of x is equal to the numerator of y , from the fact that " p " and " q " are instances of the same proposition we cannot deduce that the order of " p " is equal to that of " q ". This was shown above (p. 363) for the particular case of elementary and non-elementary propositions (Orders 0 and > 0), and obviously holds in general. Order is only a characteristic of a particular symbol which is an instance of the proposition or function.

We shall now show briefly how this theory solves the remaining contradictions of group B*.

* It may be as well to repeat that for the contradictions of group A my theory preserves the solutions given in *Principia Mathematica*.

(a) "I am lying".

This we should analyse as " $(\exists 'p', p)$; I am saying ' p '. ' p ' means $p. \sim p$ ". Here to get a definite meaning for *means** it is necessary to limit in some way the order of " p ". Suppose " p " is to be of the n th or lesser order. Then, symbolizing by ϕ_n , a function of *type* n , " p " may be $(\exists \phi_n). \phi_{n+1}(\phi_n)$.

Hence $\exists "p"$ involves $\exists \phi_{n+1}$ and "I am lying" in the sense of "I am asserting a false proposition of order n ", is at least of order $n+1$ and does not contradict itself.

(b) (1) The least integer not nameable in fewer than nineteen syllables.

(2) The least undefinable ordinal.

(3) Richard's Paradox.

All these result from the obvious ambiguity of "naming" and "defining". The name or definition is in each case a functional symbol which is only a name or definition by meaning something. The sense in which it means must be made precise by fixing its order; the name or definition involving all such names or definitions will be of a higher order, and this removes the contradiction. My solutions of these contradictions are obviously very similar to those of Whitehead and Russell, the difference between them lying merely in our different conceptions of the order of propositions and functions. For me propositions in themselves have no orders; they are just different truth-functions of atomic propositions—a definite totality, depending only on what atomic propositions there are. Orders and illegitimate totalities only come in with the symbols we use to symbolize the facts in variously complicated ways.

To sum up: in this chapter I have defined a range of predicative functions which escapes contradiction and enables us to dispense with the axiom of reducibility. And I have given a solution of the contradictions of group B which rests on and explains the fact that they all contain some epistemic element.

* When I say " p means p ", I do not suppose there to be a single object p meant by " p ". The meaning of " p " is that one of a certain set of possibilities is realized, and this meaning results from the meaning-relations of the separate signs in " p " to the real objects which it is about. It is these meaning-relations which vary with the order of " p ". And the order of " p " is limited not because p in $(\exists p)$ is limited, but by "means" which varies in meaning with the order of " p ".

IV. PROPOSITIONAL FUNCTIONS IN EXTENSION.

Before we go on, let us look round and see where we have got to. We have seen that the introduction of the notion of a predicative function has given us a range for ϕ which enables us to dispense with the axiom of reducibility. Hence it removes the second and most important defect in the theory of *Principia Mathematica*; but how do we now stand with regard to the other two difficulties, the difficulty of including all classes and relations in extension and not merely definable ones, and the difficulty connected with identity?

The difficulty about identity we can get rid of, at the cost of great inconvenience, by adopting Wittgenstein's convention, which enables us to eliminate "=" from any proposition in which it occurs. But this puts us in a hopeless position as regards classes, because, having eliminated "=" altogether, we can no longer use $x = y$ as a propositional function in defining finite classes. So that the only classes with which we are now able to deal are those defined by predicative functions.

It may be useful here to repeat the definition of a predicative function of individuals; it is any truth-function of atomic functions and atomic propositions. We call such functions "predicative" because they correspond, as nearly as a precise notion can to a vague one, to the idea that ϕa predicates the same thing of a as ϕb does of b . They include all the propositional functions which occur in *Principia Mathematica*, including identity as there defined. It is obvious, however, that we ought not to define identity in this way as agreement in respect of all predicative functions, because two things can clearly agree as regards all atomic functions and therefore as regards all predicative functions, and yet they are two things and not, as the proposed definition of identity would involve, one thing.

Hence our theory is every bit as inadequate as *Principia Mathematica* to provide an extensional logic; in fact, if we reject this false definition of identity, we are unable to include among the classes dealt with even all finite enumerated classes. Mathematics then becomes hopeless because we cannot be sure that there is any class defined by a predicative function whose number is two; for things may all fall into triads which agree in every respect, in which case there would be in our system no unit classes and no two-member classes.

If we are to preserve at all the ordinary form of mathematics, it looks as if some extension must be made in the notion of a propositional function, so as to take in other classes as well. Such an extension is desirable on other grounds, because many things which would naturally

be regarded as propositional functions can be shown not to be predicative functions.

For example

$$F(x, y) = \text{Something other than } x \text{ and } y \text{ satisfies } \hat{\phi}z.$$

(Here, of course, "other than" is to be taken strictly, and not in the *Principia Mathematica* sense of "distinguishable from".)

This is not a predicative function, but is made up of parts of two predicative functions :

(1) For $x \neq y$

$$\begin{aligned} F(x, y) \text{ is } & \phi x . \phi y . \supset . Nc' \hat{z}(\phi z) \geq 3 : . \\ & \phi x . \sim \phi y . \vee . \phi y . \sim \phi x : \supset : Nc' \hat{z}(\phi z) \geq 2 : . \\ & \sim \phi x . \sim \phi y : \supset : Nc' \hat{z}(\phi z) \geq 1 . \end{aligned}$$

This is a predicative function because it is a truth function of ϕx , ϕy and the constant propositions $Nc' \hat{z}(\phi z) \geq 1, 2, 3$, which do not involve x, y .

(2) For $x = y$

$$\begin{aligned} F(x, x) \text{ is } & \phi(x) . \supset . Nc' \hat{z}(\phi z) \geq 2 : \\ & \sim \phi x . \supset . Nc' \hat{z}(\phi z) \geq 1 , \end{aligned}$$

which is a predicative function.

But $F(x, y)$ is not itself a predicative function ; this is perhaps more difficult to see. But it is easy to see that all functions of this kind cannot be predicative, because if they were we could find a predicative function satisfied by any given individual a alone, which we clearly cannot in general do.

For suppose fa (if not, take $\sim f\hat{x}$).

$$\text{Let } a = \hat{x}(fx),$$

$$\beta = a - (a).$$

Then $\phi x =$ "There is nothing which satisfies fx except x , and members of β " applies to a and a alone. So such functions cannot always be predicative.

Just as $F(x, y)$ above, so also " $x = y$ " is made up of two predicative functions :

(1) For $x \neq y$

" $x = y$ " may be taken to be $(\exists \phi) . \phi x . \sim \phi x : (\exists \phi) . \phi y . \sim \phi y$,
i.e. a contradiction.

(2) For $x = y$

" $x = y$ " may be taken to be $(\phi) : . \phi x . v . \sim \phi x : \phi y . v . \sim \phi y$,
i.e. a tautology.

But " $x = y$ " is not itself predicative.

It seems, therefore, that we need to introduce non-predicative propositional functions. How is this to be done? The only practicable way is to do it as radically and drastically as possible; to drop altogether the notion that ϕa says about a , what ϕb says about b ; to treat propositional functions like mathematical functions, that is, extensionalize them completely. Indeed, it is clear that, mathematical functions being derived from propositional, we shall get an adequately extensional account of the former only by taking a completely extensional view of the latter.

So in addition to the previously defined concept of a predicative function, which we shall still require for certain purposes, we define, or rather explain, for in our system it must be taken as indefinable, the new concept of a propositional function in extension. Such a function of one individual results from any one-many relation in extension between propositions and individuals; that is to say a correlation, practicable or impracticable, which to every individual associates a unique proposition, the individual being the argument to the function, the proposition its value.

Thus ϕ (Socrates) may be Queen Anne is dead.

ϕ (Plato) ,, Einstein is a great man.

$\phi \hat{x}$ being simply an arbitrary association of propositions ϕx to individuals x .

A function in extension will be marked by a suffix e thus $\phi_e \hat{x}$.

Then we can talk of the totality of such functions as the range of values of an apparent variable ϕ_e .

Consider now $(\phi_e) : \phi_e x \equiv \phi_e y$.

This asserts that in any such correlation the proposition correlated with x is equivalent to that correlated with y .

If $x = y$ this is a tautology (it is the logical product of values of $p \equiv p$).

But if $x \neq y$ it is a contradiction. For in one of the correlations some p will be associated with x , and $\sim p$ with y .

Then for this correlation $f_e \hat{x}$, $f_e x$ is p , $f_e y$ is $\sim p$, so that $f_e x \equiv f_e y$ is self-contradictory and $(\phi_e) \cdot \phi_e x \equiv \phi_e y$ is self-contradictory.

So $(\phi_e) \cdot \phi_e x \equiv \phi_e y$ is a tautology if $x = y$, a contradiction if $x \neq y^*$.

Hence it can suitably be taken as the definition of $x = y$.

$x = y$ is a function in extension of two variables. Its value is tautology when x and y have the same value, contradictions when x, y have different values.

We have now to defend this suggested range of functions for a variable ϕ_e against the charges that it is illegitimate or leads to contradictions. It is legitimate because it is an intelligible notation, giving a definite meaning to the symbols in which it is employed. Nor can it lead to contradictions, for it will escape all the suggested contradictions just as the range of predicative functions will. Any symbol containing the variable ϕ_e will mean in a different way from a symbol not containing it, and we shall have the same sort of ambiguity of "meaning" as in Chapter III, which will remove the contradictions. Nor can any of the first group of contradictions be restored by our new notation, for it will still be impossible for a class to be a member of itself, as our functions in extension are confined to definite types of arguments by definition.

We have now to take the two notions we have defined, predicative functions and functions in extension, and consider when we shall want to use one and when the other†. First let us take the case when the arguments are individuals: then there is every advantage in taking the range of functions we use in mathematics to be that of functions in extension. We have seen how this enables us to define identity satisfactorily, and it is obvious that we shall need no axiom of reducibility, for any propositional function obtained by generalization, or in any manner whatever, is a function in extension. Further it will give us a satisfactory theory of classes, for any class will be defined by a function in extension, *e.g.* by the function which is tautology for any member of the class as argument, but contradiction for any other argument, and the null-class will be defined by the self-contradictory function. So the totality of classes can be reduced to that of functions in extension.

* On the other hand $(\phi) \cdot \phi x \equiv \phi y$ (ϕ predicative) is a tautology if $x = y$, but not a contradiction if $x \neq y$.

† Of course, predicative functions are also functions in extension; the question is which range we want for our variable function.

and therefore it will be this totality which we shall require in mathematics, not the totality of predicative functions, which corresponds not to "all classes" but to "all predicates" or "all properties"

On the other hand, when we get to functions of functions the situation is rather different. There appears to be no point in considering any except predicative functions of functions; the reasons for introducing functions in extension no longer apply. For we do not need to define identity between functions, but only identity between classes which reduces to equivalence between functions, which is easily defined. Nor do we wish to consider classes of functions, but classes of classes, of which a simpler treatment is also possible. So in the case of functions of functions we confine ourselves to such as are predicative.

Let us recall the definition of a predicative function of functions; it is a truth-function of their values and constant propositions*. All functions of functions which occur in *Principia* are of this sort, but "I believe (x) . ϕx " as a function of $\hat{\phi x}$ is not. Predicative functions of functions are extensional in the sense of *Principia*, that is if the range of $f(\hat{\phi x})$ be that of predicative functions of functions,

$$\phi_e x \equiv_r \psi_e x : \supset : f(\phi_e \hat{x}) \equiv f(\psi_e \hat{x}).$$

This is because $f(\phi_e \hat{x})$ is a truth-function of the values of $\phi_e x$ which are equivalent to the corresponding values of $\psi_e x$ so that $f(\phi_e \hat{x})$ is equivalent to $f(\psi_e \hat{x})$.

If we assumed this we should have a very simple theory of classes, since there would be no need to distinguish $\hat{x}(\phi_e x)$ from $\phi_e \hat{x}$. But though it is a tautology there is clearly no way of proving it, so that we should have to take it as a primitive proposition. If we wish to avoid this we have only to keep the theory of classes given in *Principia* based on "the derived extensional function". The range of predicative functions of functions is adequate to deal with classes of classes because, although, as we have seen, there may be classes of individuals which can only be

* It is, I think, predicative functions of functions which Mr. Russell in the introduction to the second edition of *Principia* tries to describe as functions into which functions enter only through their values. But this is clearly an insufficient description, because $\hat{\phi x}$ only enters into $F(\hat{\phi x}) = "I \text{ believe } \phi a"$ through its value ϕa , but this is certainly not a function of the kind meant, for it is not extensional. I think the point can only be explained by introducing, as I have, the notion of a truth function. To contend, as Mr. Russell does, that all functions of functions are predicative is to embark on a futile verbal dispute, owing to the ambiguity of the vague term functions of functions which may be used to mean only such as are predicative or to include also such as $F(\hat{\phi x})$ above.

defined by functions in extension, yet any class of classes can be defined by a predicative function, namely by $f(\alpha)$ where

$$f(\hat{\phi}_e x) = \Sigma_{\psi} (\phi_e x \equiv_x \psi_e x),$$

i. e. the logical sum of $\phi_e x \equiv_x \psi_e x$ for all the functions $\psi_e \hat{x}$ which define the members of the class of classes. Of course, if the class of classes is infinite, this expression cannot be written down. But, nevertheless, there will be the logical sum of these functions, though we cannot express it*.

So to obtain a complete theory of classes we must take the range of functions of individuals to be that of functions in extension; but the range of functions of functions to be that of predicative functions. By using these variables we obtain the system of *Principia Mathematica*, simplified by the omission of the axiom of reducibility, and a few corresponding alterations. Formally it is almost unaltered; but its meaning has been considerably changed. And in thus preserving the form while modifying the interpretation I am following the great school of mathematical logicians, who, in virtue of a series of startling definitions, have saved mathematics from the sceptics, and provided a rigid demonstration of its propositions. Only so can we preserve it from the Bolshevik menace of Brouwer and Weyl.

V. THE AXIOMS.

I have shown in the last two chapters how to remedy the three principal defects in *Principia Mathematica* as a foundation for mathematics. Now we have to consider the two important difficulties which remain, which concern the Axiom of Infinity and the Multiplicative Axiom. The introduction of these two axioms is not so grave as that of the Axiom of Reducibility, because they are not in themselves such objectionable assumptions, and because mathematics is largely independent of the Multiplicative Axiom, and might reasonably be supposed to require an axiom of infinity. Nevertheless, we must try to determine the logical status of these axioms, whether they are tautologies or empirical propositions or even contradictions. In this inquiry I shall

* A logical sum is not like an algebraic sum; only a finite number of terms can have an algebraic sum, for an "infinite sum" is really a limit. But the logical sum of a set of propositions is the proposition that these are not all false, and exists whether the set be finite or infinite.

include, from curiosity, the Axiom of Reducibility, although, since we have dispensed with it, it no longer really concerns us.

Let us begin with the Axiom of Reducibility, which asserts that all functions of individuals obtained by the generalization of matrices are equivalent to elementary functions. In discussing it several cases arise, of which I shall consider only the most interesting, that, namely, in which the numbers of individuals and of atomic functions of individuals are both infinite. In this case the axiom is an empirical proposition, that is to say, neither a tautology nor a contradiction, and can therefore be neither asserted nor denied by logic or mathematics. This is shown as follows :—

(a) The axiom is not a contradiction, but may be true.

For it is clearly possible that there should be an atomic function defining every class of individuals. In which case every function would be equivalent not merely to an elementary but to an atomic function.

(b) The axiom is not a tautology, but may be false.

For it is clearly possible that there should be an infinity of atomic functions, and an individual a such that whichever atomic function we take there is another individual agreeing with a in respect of all the other functions, but not in respect of the function taken. Then $(\phi). \phi! x \equiv \phi! a$ could not be equivalent to any elementary function of x .

Having thus shown that the Axiom of Reducibility is neither a tautology nor a contradiction, let us proceed to the Multiplicative Axiom. This asserts that, given any existent class K of existent classes, there is a class having exactly one member in common with each member of K . If by "class" we mean, as I do, any set of things homogeneous in type not necessarily definable by a function which is not merely a function in extension, the Multiplicative Axiom seems to me the most evident tautology. I cannot see how this can be the subject of reasonable doubt, and I think it never would have been doubted unless it had been misinterpreted. For with the meaning it has in *Principia*, where the class whose existence it asserts must be one definable by a propositional function of the sort which occurs in *Principia*, it becomes really doubtful and, like the Axiom of Reducibility, neither a tautology nor a contradiction. We prove this by showing

(a) It is not a contradiction.

For it is clearly possible that every class (in my sense) should be defined by an atomic function, so that, since there is bound to be a

class in my sense having one member in common with each member of K , this would be also a class in the sense of *Principia*.

(b) It is not a tautology.

To show this we take not the Multiplicative Axiom itself but the equivalent theorem that any two classes are commensurable.

Consider then the following case: let there be no atomic functions of two or more variables, and only the following atomic functions of one variable:—

Associated with each individual a an atomic function $\phi_a \hat{x}$ such that

$$\phi_a x . \equiv_x x = a.$$

One other atomic function $f \hat{x}$ such that $\hat{x}(fx)$, $\hat{x}(\sim fx)$ are both infinite classes.

Then there is no one-one relation, in the sense of *Principia*, having either $\hat{x}(fx)$ or $\hat{x}(\sim fx)$ for domain, and therefore these two classes are incommensurable.

Hence the multiplicative axiom, interpreted as it is in *Principia*, is not a tautology but logically doubtful. But, as I interpret it, it is an obvious tautology, and this can be claimed as an additional advantage in my theory. It will probably be objected that, if it is a tautology, it ought to be able to be proved, *i.e.* deduced from the simpler primitive propositions which suffice for the deduction of the rest of mathematics. But it does not seem to me in the least unlikely that there should be a tautology, which could be stated in finite terms, whose proof was, nevertheless, infinitely complicated and therefore impossible for us. Moreover, we cannot expect to prove the multiplicative axiom in my system, because my system is formally the same as that of *Principia*, and the multiplicative axiom obviously cannot be proved in the system of *Principia*, in which it is not a tautology.

We come now to the Axiom of Infinity, of which again my system and that of *Principia* give different interpretations. In *Principia*, owing to the definition of identity there used, the axiom means that there are an infinity of distinguishable individuals, which is an empirical proposition; since, even supposing there to be an infinity of individuals, logic cannot determine whether there are an infinity of them no two of which have all their properties in common; but on my system, which admits functions in extension, the axiom of infinity asserts merely that there is an infinite number of individuals. This appears equally to be a mere question of fact; but the profound analysis of Wittgenstein has shown

that this is an illusion, and that, if it means anything, it must be either a tautology or contradiction. This will be much easier to explain if we begin not with infinity but with some smaller number.

Let us start with "There is an individual", or writing it as simply as possible in logical notation,

$$"(\exists x). x = x"$$

Now what is this proposition? It is the logical sum of the tautologies $x = x$ for all values of x , and is therefore a tautology. But suppose there were no individuals, and therefore no values of x , then the above formula is absolute nonsense. So, if it means anything, it must be a tautology.

Next let us take "There are at least two individuals, or

$$"(\exists x, y). x \neq y"$$

This is the logical sum of the propositions $x \neq y$, which are tautologies if x and y have different values, contradictions if they have the same value. Hence it is the logical sum of a set of tautologies and contradictions; and therefore a tautology, if any one of the set is a tautology, but otherwise a contradiction. That is, it is a tautology if x and y can take different values (*i.e.* if there are two individuals), but otherwise a contradiction.

A little reflection will make it clear that this will hold not merely of 2, but of any other number, finite or infinite. That is, "There are at least n individuals" is always either a tautology or contradiction, never a genuine proposition. We cannot, therefore, say anything about the number of individuals, since, when we attempt to do so, we never succeed in constructing a genuine proposition, but only a formula which is either tautological or self-contradictory. The number of individuals can, in Wittgenstein's phrase, only be shown, and it will be shown by whether the above formulae are tautological or contradictory.

The sequence "There is an individual".

"There are at least 2 individuals".

"There are at least n individuals".

"There are at least \aleph_0 individuals".

" " " \aleph_1 "

begins by being tautologous; but somewhere it begins to be contradictory,

and the position of the last tautologous term shows the number of individuals.

It may be wondered how, if we can say nothing about it, we can envisage as distinct possibilities that the number of individuals in the world is so-and-so. We do this by imagining different universes of discourse, to which we may be confined, so that by "all" we mean all in the universe of discourse and then that such and such a universe contains so and so many individuals is a real possibility, and can be asserted in a genuine proposition. It is only when we take, not a limited universe of discourse, but the whole world, that nothing can be said about the number of individuals in it.

We can do logic not only for the whole world but also for such limited universes of discourse; if we take one containing n individuals,

$Nc'\hat{x}(x = x) \geq n$ will be a tautology,

$Nc'\hat{x}(x = x) \geq n+1$ a contradiction.

Hence $Nc'\hat{x}(x = x) \geq n+1$ cannot be deduced from the primitive propositions common to all universes, and therefore for a universe containing $n+1$ individuals must be taken as a primitive proposition.

Similarly the axiom of infinity in the logic of the whole world, if it is a tautology, cannot be proved, but must be taken as a primitive proposition. And this is the course which we must adopt, unless we prefer the view that all analysis is self-contradictory and meaningless. We do not have to assume that any particular set of things, *e.g.* atoms, is infinite, but merely that there is some infinite type which we can take to be the type of individuals.