# *Mathematical Logic as based on the Theory of Types.*

BY BERTRAND RUSSELL.

The following theory of symbolic logic recommended itself to me in the first instance by its ability to solve certain contradictions, of which the one best known to mathematicians is Burali-Forti's concerning the greatest ordinal.* But the theory in question seems not wholly dependent on this indirect recommendation; it has also, if I am not mistaken, a certain consonance with common sense which makes it inherently credible. This, however, is not a merit upon which much stress should be laid; for common sense is far more fallible than it likes to believe. I shall therefore begin by stating some of the contradictions to be solved, and shall then show how the theory of logical types effects their solution.

## I.

### *The Contradictions.*

(1) The oldest contradiction of the kind in question is the *Epimenides*. Epimenides the Cretan said that all Cretans were liars, and all other statements made by Cretans were certainly lies. Was this a lie? The simplest form of this contradiction is afforded by the man who says "I am lying;" if he is lying, he is speaking the truth, and vice versa.

(2) Let $w$ be the class of all those classes which are not members of themselves. Then, whatever class $x$ may be, "$x$ is a $w$" is equivalent† to "$x$ is not an $x$." Hence, giving to $x$ the value $w$, "$w$ is a $w$" is equivalent to "$w$ is not a $w$."

(3) Let $T$ be the relation which subsists between two relations $R$ and $S$ whenever $R$ does not have the relation $R$ to $S$. Then, whatever relations $R$ and $S$ may be, "$R$ has the relation $T$ to $S$" is equivalent to "$R$ does not have the

---

*See below.

† Two propositions are called *equivalent* when both are true or both are false.

relation $R$ to $S$." Hence, giving the value $T$ to both $R$ and $S$, "$T$ has the relation $T$ to $T$" is equivalent to "$T$ does not have the relation $T$ to $T$."

(4) The number of syllables in the English names of finite integers tends to increase as the integers grow larger, and must gradually increase indefinitely, since only a finite number of names can be made with a given finite number of syllables. Hence the names of some integers must consist of at least nineteen syllables, and among these there must be a least. Hence "the least integer not nameable in fewer than nineteen syllables" must denote a definite integer; in fact, it denotes 111,777. But "the least integer not nameable in fewer than nineteen syllables" is itself a name consisting of eighteen syllables; hence the least integer not nameable in fewer than nineteen syllables can be named in eighteen syllables, which is a contradiction.*

(5) Among transfinite ordinals some can be defined, while others can not; for the total number of possible definitions is $\aleph_0$, while the number of transfinite ordinals exceeds $\aleph_0$. Hence there must be indefinable ordinals, and among these there must be a least. But this is defined as "the least indefinable ordinal," which is a contradiction.†

(6) Richard's paradox ‡ is akin to that of the least indefinable ordinal. It is as follows: Consider all decimals that can be defined by means of a finite number of words; let $E$ be the class of such decimals. Then $E$ has $\aleph_0$ terms; hence its members can be ordered as the 1st, 2nd, 3rd, . . . . . Let $N$ be a number defined as follows: If the $n$th figure in the $n$th decimal is $p$, let the $n$th figure in $N$ be $p+1$ (or 0, if $p=9$). Then $N$ is different from all the members of $E$, since, whatever finite value $n$ may have, the $n$th figure in $N$ is different from the $n$th figure in the $n$th of the decimals composing $E$, and therefore $N$ is different from the $n$th decimal. Nevertheless we have defined $N$ in a finite number of words, and therefore $N$ ought to be a member of $E$. Thus $N$ both is and is not a member of $E$.

(7) Burali-Forti's contradiction § may be stated as follows: It can be shown

---

* This contradiction was suggested to me by Mr. G. G. Berry of the Bodleian Library.

† *Cf.* König, "Ueber die Grundlagen der Mengenlehre und das Kontinuumproblem," *Math. Annalen*, Vol. LXI (1905); A. C. Dixon, "On 'well-ordered' aggregates," *Proc. London Math. Soc.*, Series 2, Vol. IV, Part I (1906); and E. W. Hobson, "On the Arithmetic Continuum," *ibid.* The solution offered in the last of these papers does not seem to me adequate.

‡ *Cf.* Poincaré, "Les mathématiques et la logique," *Revue de Métaphysique et de Morale*, Mai, 1906, especially sections VII and IX; also Peano, *Revista de Mathematica*, Vol. VIII, No. 5 (1906), p. 149 ff.

§ "Una questione sui numeri transfiniti," *Rendiconti del circolo matematico di Palermo*, Vol. XI (1897).

that every well-ordered series has an ordinal number, that the series of ordinals up to and including any given ordinal exceeds the given ordinal by one, and (on certain very natural assumptions) that the series of all ordinals (in order of magnitude) is well-ordered. It follows that the series of all ordinals has an ordinal number, $\Omega$ say. But in that case the series of all ordinals including $\Omega$ has the ordinal number $\Omega + 1$, which must be greater than $\Omega$. Hence $\Omega$ is not the ordinal number of all ordinals.

In all the above contradictions (which are merely selections from an indefinite number) there is a common characteristic, which we may describe as self-reference or reflexiveness. The remark of Epimenides must include itself in its own scope If *all* classes, provided they are not members of themselves, are members of $w$, this must also apply to $w$; and similarly for the analogous relational contradiction. In the cases of names and definitions, the paradoxes result from considering non-nameability and indefinability as elements in names and definitions. In the case of Burali-Forti's paradox, the series whose ordinal number causes the difficulty is the series of all ordinal numbers. In each contradiction something is said about *all* cases of some kind, and from what is said a new case seems to be generated, which both is and is not of the same kind as the cases of which *all* were concerned in what was said. Let us go through the contradictions one by one and see how this occurs.

(1) When a man says "I am lying," we may interpret his statement as: "There is a proposition which I am affirming and which is false." All statements that "there is" so-and-so may be regarded as denying that the opposite is always true; thus "I am lying" becomes: "It is not true of all propositions that either I am not affirming them or they are true;" in other words, "It is not true for all propositions $p$ that if I affirm $p$, $p$ is true." The paradox results from regarding this statement as affirming a proposition, which must therefore come within the scope of the statement. This, however, makes it evident that the notion of "all propositions" is illegitimate; for otherwise, there must be propositions (such as the above) which are about all propositions, and yet can not, without contradiction, be included among the propositions they are about. Whatever we suppose to be the totality of propositions, statements about this totality generate new propositions which, on pain of contradiction, must lie outside the totality. It is useless to enlarge the totality, for that equally enlarges the scope of statements about the totality. Hence there must be no totality of propositions, and "all propositions" must be a meaningless phrase.

(2) In this case, the class $w$ is defined by reference to "all classes," and then turns out to be one among classes. If we seek help by deciding that no class is a member of itself, then $w$ becomes the class of all classes, and we have to decide that this is not a member of itself, *i. e.*, is not a class. This is only possible if there is no such thing as the class of all classes in the sense required by the paradox. That there is no such class results from the fact that, if we suppose there is, the supposition immediately gives rise (as in the above contradiction) to new classes lying outside the supposed total of all classes.

(3) This case is exactly analogous to (2), and shows that we can not legitimately speak of "all relations."

(4) "The least integer not nameable in fewer than nineteen syllables" involves the totality of names, for it is "the least integer such that all names either do not apply to it or have more than nineteen syllables." Here we assume, in obtaining the contradiction, that a phrase containing "all names" is itself a name, though it appears from the contradiction that it can not be one of the names which were supposed to be all the names there are. Hence "all names" is an illegitimate notion.

(5) This case, similarly, shows that "all definitions" is an illegitimate notion.

(6) This is solved, like (5), by remarking that "all definitions" is an illegitimate notion. Thus the number $E$ is *not* defined in a finite number of words, being in fact not defined at all.*

(7) Burali-Forti's contradiction shows that "all ordinals" is an illegitimate notion; for if not, all ordinals in order of magnitude form a well-ordered series, which must have an ordinal number greater than all ordinals.

Thus all our contradictions have in common the assumption of a totality such that, if it were legitimate, it would at once be enlarged by new members defined in terms of itself.

This leads us to the rule: "Whatever involves *all* of a collection must not be one of the collection;" or, conversely: "If, provided a certain collection had a total, it would have members only definable in terms of that total, then the said collection has no total."†

---

* *Cf.* "Les paradoxes de la logique," by the present author, *Revue de Métaphysique et de Morale*, Sept., 1906, p. 645.

† When I say that a collection has no total, I mean that statements about *all* its members are nonsense. Furthermore, it will be found that the use of this principle requires the distinction of *all* and *any* considered in Section II.

The above principle is, however, purely negative in its scope. It suffices to show that many theories are wrong, but it does not show how the errors are to be rectified. We can not say: "When I speak of *all* propositions, I mean all except those in which 'all propositions' are mentioned;" for in this explanation we have mentioned the propositions in which all propositions are mentioned, which we can not do significantly. It is impossible to avoid mentioning a thing by mentioning that we won't mention it. One might as well, in talking to a man with a long nose, say: "When I speak of noses, I except such as are inordinately long," which would not be a very successful effort to avoid a painful topic. Thus it is necessary, if we are not to sin against the above negative principle, to construct our logic without mentioning such things as "all propositions" or "all properties," and without even having to say that we are excluding such things The exclusion must result naturally and inevitably from our positive doctrines, which must make it plain that "all propositions" and "all properties" are meaningless phrases.

The first difficulty that confronts us is as to the fundamental principles of logic known under the quaint name of "laws of thought." "All propositions are either true or false," for example, has become meaningless. If it were significant, it would be a proposition, and would come under its own scope. Nevertheless, some substitute must be found, or all general accounts of deduction become impossible.

Another more special difficulty is illustrated by the particular case of mathematical induction. We want to be able to say: "If $n$ is a finite integer, $n$ has all properties possessed by 0 and by the successors of all numbers possessing them." But here "all properties" must be replaced by some other phrase not open to the same objections. It might be thought that "all properties possessed by 0 and by the successors of all numbers possessing them" might be legitimate even if "all properties" were not. But in fact this is not so. We shall find that phrases of the form "all properties which *etc.*" involve *all* properties of which the "*etc.*" can be significantly either affirmed or denied, and not only those which in fact have whatever characteristic is in question; for, in the absence of a catalogue of properties having this characteristic, a statement about all those that have the characteristic must be hypothetical, and of the form: "It is always true that, if a property has the said characteristic, then etc." Thus mathematical induction is *primâ facie* incapable of being significantly enunciated, if "all properties" is a phrase destitute of meaning. This difficulty, as we shall see later, can be avoided; for the present we must consider the laws of logic, since these are far more fundamental.

## II.

### *All and Any.*

Given a statement containing a variable $x$, say "$x = x$," we may affirm that this holds in all instances, or we may affirm any one of the instances without deciding as to which instance we are affirming. The distinction is roughly the same as that between the general and particular enunciation in Euclid. The general enunciation tells us something about (say) all triangles, while the particular enunciation takes one triangle, and asserts the same thing of this one triangle. But the triangle taken is *any* triangle, not some one special triangle; and thus although, throughout the proof, only one triangle is dealt with, yet the proof retains its generality. If we say: "Let $ABC$ be a triangle, then the sides $AB$, $AC$ are together greater than the side $BC$," we are saying something about *one* triangle, not about *all* triangles; but the one triangle concerned is absolutely ambiguous, and our statement consequently is also absolutely ambiguous. We do not affirm any one definite proposition, but an undetermined one of all the propositions resulting from supposing $ABC$ to be this or that triangle. This notion of ambiguous assertion is very important, and it is vital not to confound an ambiguous assertion with the definite assertion that the same thing holds in *all* cases.

The distinction between (1) asserting any value of a propositional function, and (2) asserting that the function is always true, is present throughout mathematics, as it is in Euclid's distinction of general and particular enunciations. In any chain of mathematical reasoning, the objects whose properties are being investigated are the arguments to *any* value of some propositional function. Take as an illustration the following definition:

"We call $f(x)$ continuous for $x = a$ if, for every positive number $\sigma$, different from 0, there exists a positive number $\varepsilon$, different from 0, such that, for all values of $\delta$ which are numerically less than $\varepsilon$, the difference $f(a + \delta) - f(a)$ is numerically less than $\sigma$."

Here the function $f$ is *any* function for which the above statement has a meaning; the statement is *about* $f$, and varies as $f$ varies. But the statement is not *about* $\sigma$ or $\varepsilon$ or $\delta$, because *all* possible values of these are concerned, not one undetermined value. (In regard to $\varepsilon$, the statement "there exists a positive number $\varepsilon$ such that *etc.*" is the denial that the denial of "*etc.*" is true of *all* positive numbers.) For this reason, when *any* value of a propositional function is asserted, the argument (*e.g.*, $f$ in the above) is called a *real* variable; whereas, when a function is said to be *always* true, or to be not always true, the argument is called an *apparent* variable.* Thus in the above definition, $f$ is a real variable, and $\sigma$, $\varepsilon$, $\delta$ are apparent variables.

When we assert *any* value of a propositional function, we shall say simply that we assert the *propositional function.* Thus if we enunciate the law of identity in the form "$x = x$," we are asserting the function "$x = x$;" *i. e.*, we are asserting any value of this function. Similarly we may be said to deny a propositional function when we deny any instance of it. We can only truly assert a propositional function if, whatever value we choose, that value is true; similarly we can only truly deny it if, whatever value we choose, that value is false. Hence in the general case, in which some values are true and some false, we can neither assert nor deny a propositional function.†

If $\phi x$ is a propositional function, we will denote by "$(x) . \phi x$" the proposition "$\phi x$ is always true." Similarly "$(x, y) . \phi(x, y)$" will mean "$\phi(x, y)$ is always true," and so on. Then the distinction between the assertion of all values and the assertion of any is the distinction between (1) asserting $(x) . \phi x$ and (2) asserting $\phi x$ where $x$ is undetermined. The latter differs from the former in that it can not be treated as one determinate proposition.

The distinction between asserting $\phi x$ and asserting $(x) . \phi x$ was, I believe, first emphasized by Frege.‡ His reason for introducing the distinction explicitly was the same which had caused it to be present in the practice of mathematicians; namely, that deduction can only be effected with *real* variables, not with apparent variables. In the case of Euclid's proofs, this is evident: we need (say) some one triangle $ABC$ to reason about, though it does not matter what triangle it is. The triangle $ABC$ is a *real* variable; and although it is *any* triangle, it remains the *same* triangle throughout the argument. But in the general enunciation,

---

* These two terms are due to Peano, who uses them approximately in the above sense. *Cf.*, *e. g.*, *Formulaire Mathématique*, Vol. IV, p. 5 (Turin, 1903).

† Mr. MacColl speaks of "propositions" as divided into the three classes of certain, variable, and impossible. We may accept this division as applying to propositional functions. A function which can be asserted is certain, one which can be denied is impossible, and all others are (in Mr. MacColl's sense) variable.

‡ See his *Grundgesetze der Arithmetik*, Vol. I (Jena, 1893), § 17, p. 31.

the triangle is an apparent variable. If we adhere to the apparent variable, we can not perform any deductions, and this is why in all proofs, real variables have to be used. Suppose, to take the simplest case, that we know "$\phi x$ is always true," *i. e.* "$(x) . \phi x$," and we know "$\phi x$ always implies $\psi x$," *i. e.* "$(x) . \{\phi x$ implies $\psi x\}$." How shall we infer "$\psi x$ is always true," *i. e.* "$(x) . \psi x$?" We know it is always true that if $\phi x$ is true, and if $\phi x$ implies $\psi x$, then $\psi x$ is true. But we have no premises to the effect that $\phi x$ is true and $\phi x$ implies $\psi x$; what we have is: $\phi x$ is *always* true, and $\phi x$ *always* implies $\psi x$. In order to make our inference, we must go from "$\phi x$ is always true" to $\phi x$, and from "$\phi x$ always implies $\psi x$" to "$\phi x$ implies $\psi x$," where the $x$, while remaining any possible argument, is to be the same in both. Then, from "$\phi x$" and "$\phi x$ implies $\psi x$," we infer "$\psi x$;" thus $\psi x$ is true for any possible argument, and therefore is always true. Thus in order to infer "$(x) . \psi x$ from "$(x) . \phi x$" and "$(x) . \{\phi x$ implies $\psi x\}$," we have to pass from the apparent to the real variable, and then back again to the apparent variable. This process is required in all mathematical reasoning which proceeds from the assertion of all values of one or more propositional functions to the assertion of all values of some other propositional function, as, *e. g.*, from "all isosceles triangles have equal angles at the base" to "all triangles having equal angles at the base are isosceles." In particular, this process is required in proving *Barbara* and the other moods of the syllogism. In a word, *all deduction operates with real variables* (or with constants).

It might be supposed that we could dispense with apparent variables altogether, contenting ourselves with *any* as a substitute for *all*. This, however, is not the case. Take, for example, the definition of a continuous function quoted above: in this definition $\sigma$, $\varepsilon$, and $\delta$ must be apparent variables. Apparent variables are constantly required for definitions. Take, *e. g.*, the following: "An integer is called a *prime* when it has no integral factors except 1 and itself." This definition unavoidably involves an apparent variable in the form: "If $n$ is an integer other than 1 or the given integer, $n$ is not a factor of the given integer, for all possible values of $n$"

The distinction between *all* and *any* is, therefore, necessary to deductive reasoning, and occurs throughout mathematics; though, so far as I know, its importance remained unnoticed until Frege pointed it out.

For our purposes it has a different utility, which is very great. In the case of such variables as propositions or properties, "any value" is legitimate, though "all values" is not. Thus we may say: "$p$ is true or false, where $p$ is any proposition," though we can not say "all propositions are true or false." The reason is that, in the former, we merely affirm an undetermined one of the propositions of the form "$p$ is true or false," whereas in the latter we affirm (if anything) a new proposition, different from all the propositions of the form "$p$ is true or false." Thus we may admit "any value" of a variable in cases where "all values" would lead to reflexive fallacies; for the admission of "any value" does not in the same way create new values. Hence the fundamental laws of logic can be stated concerning *any* proposition, though we can not significantly say that they hold of *all* propositions. These laws have, so to speak, a particular enunciation but no general enunciation. There is no one proposition which *is* the law of contradiction (say); there are only the various instances of the law. Of any proposition $p$, we can say: "$p$ and not-$p$ can not both be true;" but there is no such proposition as: "Every proposition $p$ is such that $p$ and not-$p$ can not both be true."

A similar explanation applies to properties. We can speak of *any* property of $x$, but not of *all* properties, because new properties would be thereby generated. Thus we can say: "If $n$ is a finite integer, and if 0 has the property $\phi$, and $m+1$ has the property $\phi$ provided $m$ has it, it follows that $n$ has the property $\phi$." Here we need not specify $\phi$; $\phi$ stands for "any property." But we can not say: "A finite integer is defined as one which has *every* property $\phi$ possessed by 0 and by the successors of possessors." For here it is essential to consider *every* property,* not *any* property; and in using such a definition we assume that it embodies a *property* distinctive of finite integers, which is just the kind of assumption from which, as we saw, the reflexive contradictions spring.

In the above instance, it is necessary to avoid the suggestions of ordinary language, which is not suitable for expressing the distinction required. The point may be illustrated further as follows: If induction is to be used for defining finite integers, induction must state a definite property of finite integers, not an ambiguous property. But if $\phi$ is a real variable, the statement "$n$ has the property $\phi$ provided this property is possessed by 0 and by the successors of possessors" assigns to $n$ a property which varies as $\phi$ varies, and such a property can not be used to define the class of finite integers. We wish to say: "'$n$ is a finite integer' means: 'Whatever property $\phi$ may be, $n$ has the property $\phi$ pro-

* This is indistinguishable from "all properties."

vided $\phi$ is possessed by 0 and by the successors of possessors.'" But here $\phi$ has become an *apparent* variable. To keep it a real variable, we should have to say: "Whatever property $\phi$ may be, '$n$ is a finite integer' means: '$n$ has the property $\phi$ provided $\phi$ is possessed by 0 and by the successors of possessors.'" But here the meaning of '$n$ is a finite integer' varies as $\phi$ varies, and thus such a definition is impossible. This case illustrates an important point, namely the following: "The scope* of a real variable can never be less than the whole propositional function in the assertion of which the said variable occurs." That is, if our propositional function is (say) "$\phi x$ implies $p$," the assertion of this function will mean "any value of '$\phi x$ implies $p$' is true," *not* "'any value of $\phi x$ is true' implies $p$." In the latter, we have really "*all* values of $\phi x$ are true," and the $x$ is an *apparent* variable.

## III.

### *The Meaning and Range of Generalized Propositions.*

In this section we have to consider first the meaning of propositions in which the word *all* occurs, and then the kind of collections which admit of propositions about all their members.

It is convenient to give the name *generalized propositions* not only to such as contain *all*, but also to such as contain *some* (undefined). The proposition "$\phi x$ is sometimes true" is equivalent to the denial of "not-$\phi x$ is always true;" "some $A$ is $B$" is equivalent to the denial of "all $A$ is not $B$;" *i. e.*, of "no $A$ is $B$." Whether it is possible to find interpretations which distinguish "$\phi x$ is sometimes true" from the denial of "not-$\phi x$ is always true," it is unnecessary to inquire; for our purposes we may *define* "$\phi x$ is sometimes true" as the denial of "not-$\phi x$ is always true." In any case, the two kinds of propositions require the same kind of interpretation, and are subject to the same limitations. In each there is an apparent variable; and it is the presence of an apparent variable which constitutes what I mean by a generalized proposition. (Note that there can not be a *real* variable in any proposition; for what contains a real variable is a propositional function, not a proposition.)

The first question we have to ask in this section is: How are we to interpret the word *all* in such propositions as "all men are mortal?" At first sight, it might be thought that there could be no difficulty, that "all men" is a perfectly

---

* The *scope* of a real variable is the whole function of which "any value" is in question. Thus in "$\phi x$ implies $p$" the scope of $x$ is not $\phi x$, but "$\phi x$ implies $p$."

clear idea, and that we say of all men that they are mortal. But to this view there are many objections.

(1) If this view were right, it would seem that "all men are mortal" could not be true if there were no men. Yet, as Mr. Bradley has urged,* "Trespassers will be prosecuted" may be perfectly true even if no one trespasses; and hence, as he further argues, we are driven to interpret such propositions as hypotheticals, meaning "if anyone trespasses, he will be prosecuted;" *i. e.*, "if $x$ trespasses, $x$ will be prosecuted," where the range of values which $x$ may have, whatever it is, is certainly not confined to those who really trespass. Similarly "all men are mortal" will mean "if $x$ is a man, $x$ is mortal, where $x$ may have any value within a certain range." What this range is, remains to be determined; but in any case it is wider than "men," for the above hypothetical is certainly often true when $x$ is not a man.

(2) "All men" is a denoting phrase; and it would appear, for reasons which I have set forth elsewhere,† that denoting phrases never have any meaning in isolation, but only enter as constituents into the verbal expression of propositions which contain no constituent corresponding to the denoting phrases in question. That is to say, a denoting phrase is defined by means of the propositions in whose verbal expression it occurs. Hence it is impossible that these propositions should acquire their meaning through the denoting phrases; we must find an independent interpretation of the propositions containing such phrases, and must not use these phrases in explaining what such propositions mean. Hence we can not regard "all men are mortal" as a statement about "all men."

(3) Even if there were such an object as "all men," it is plain that it is not this object to which we attribute mortality when we say "all men are mortal." If we were attributing mortality to this object, we should have to say "*all men* is mortal." Thus the supposition that there is such an object as "all men" will not help us to interpret "all men are mortal."

(4) It seems obvious that, if we meet something which may be a man or may be an angel in disguise, it comes within the scope of "all men are mortal" to assert "if this is a man, it is mortal." Thus again, as in the case of the trespassers, it seems plain that we are really saying "if anything is a man, it is mortal," and that the question whether this or that is a man does not fall within the scope of our assertion, as it would do if the *all* really referred to "all men."

* *Logic*, Part I, Chapter II.

† "On Denoting," *Mind*, October, 1905.

(5) We thus arrive at the view that what is meant by "all men are mortal" may be more explicitly stated in some such form as "it is always true that if $x$ is a man, $x$ is mortal." Here we have to inquire as to the scope of the word *always*.

(6) It is obvious that *always* includes some cases in which $x$ is not a man, as we saw in the case of the disguised angel. If $x$ were limited to the case when $x$ is a man, we could infer that $x$ is a mortal, since if $x$ is a man, $x$ is a mortal. Hence, with the same meaning of *always*, we should find "it is always true that $x$ is mortal." But it is plain that, without altering the meaning of *always*, this new proposition is false, though the other was true.

(7) One might hope that "always" would mean "for all values of $x$." But "all values of $x$," if legitimate, would include as parts "all propositions" and "all functions," and such illegitimate totalities. Hence the values of $x$ must be somehow restricted within some legitimate totality. This seems to lead us to the traditional doctrine of a "universe of discourse" within which $x$ must be supposed to lie.

(8) Yet it is quite essential that we should have some meaning of *always* which does not have to be expressed in a restrictive hypothesis as to $x$. For suppose "always" means "whenever $x$ belongs to the class $i$." Then "all men are mortal" becomes "whenever $x$ belongs to the class $i$, if $x$ is a man, $x$ is mortal;" *i. e.*, "it is always true that if $x$ belongs to the class $i$, then, if $x$ is a man, $x$ is mortal." But what is our new *always* to mean? There seems no more reason for restricting $x$, in this new proposition, to the class $i$, than there was before for restricting it to the class *man*. Thus we shall be led on to a new wider universe, and so on *ad infinitum*, unless we can discover some natural restriction upon the possible values of (*i. e.*, some restriction given with) the function "if $x$ is a man, $x$ is mortal," and not needing to be imposed from without.

(9) It seems obvious that, since all men are mortal, there can not be any *false* proposition which is a value of the function "if $x$ is a man, $x$ is mortal." For if this is a proposition at all, the hypothesis "$x$ is a man" must be a proposition, and so must the conclusion "$x$ is mortal." But if the hypothesis is false, the hypothetical is true; and if the hypothesis is true, the hypothetical is true. Hence there can be no false propositions of the form "if $x$ is a man, $x$ is mortal."

(10) It follows that, if any values of $x$ are to be excluded, they can only be values for which there is no proposition of the form "if $x$ is a man, $x$ is mortal;"

*i.e.*, for which this phrase is meaningless. Since, as we saw in (7), there must be excluded values of $x$, it follows that the function "if $x$ is a man, $x$ is mortal" must have a certain *range of significance*,* which falls short of all imaginable values of $x$, though it exceeds the values which are men. The restriction on $x$ is therefore a restriction to the range of significance of the function "if $x$ is a man, $x$ is mortal."

(11) We thus reach the conclusion that "all men are mortal" means "if $x$ is a man, $x$ is mortal, always," where *always* means "for all values of the function 'if $x$ is a man, $x$ is mortal.'" This is an *internal* limitation upon $x$, given by the nature of the function; and it is a limitation which does not require explicit statement, since it is impossible for a function to be true more generally than for all its values. Moreover, if the range of significance of the function is $i$, the function "if $x$ is an $i$, then if $x$ is a man, $x$ is mortal" has the same range of significance, since it can not be significant unless its constituent "if $x$ is a man, $x$ is mortal" is significant. But here the range of significance is again implicit, as it was in 'if $x$ is a man, $x$ is mortal;' thus we can not make ranges of significance explicit, since the attempt to do so only gives rise to a new proposition in which the same range of significance is implicit.

Thus generally: "$(x) . \phi x$" is to mean "$\phi x$ always." This may be interpreted, though with less exactitude, as "$\phi x$ is always true," or, more explicitly: "All propositions of the form $\phi x$ are true," or "All values of the function $\phi x$ are true."† Thus the fundamental *all* is "all values of a propositional function," and every other *all* is derivative from this. And every propositional function has a certain *range of significance*, within which lie the arguments for which the function has values. Within this range of arguments, the function is true or false; outside this range, it is nonsense.

The above argumentation may be summed up as follows:

The difficulty which besets attempts to restrict the variable is, that restrictions naturally express themselves as hypotheses that the variable is of such or such a kind, and that, when so expressed, the resulting hypothetical is free from the intended restriction. For example, let us attempt to restrict the

*A function is said to be significant for the argument $x$ if it has a value for this argument. Thus we may say shortly "$\phi x$ is significant," meaning "the function $\phi$ has a value for the argument $x$." The range of significance of a function consists of all the arguments for which the function is true, together with all the arguments for which it is false.

†A linguistically convenient expression for this idea is: "$\phi x$ is true for all *possible* values of $x$," a possible value being understood to be one for which $\phi x$ is significant.

variable to *men*, and assert that, subject to this restriction, "$x$ is mortal" is always true. Then what is always true is that if $x$ is a man, $x$ is mortal; and this hypothetical is true even when $x$ is not a man. Thus a variable can never be restricted within a certain range if the propositional function in which the variable occurs remains significant when the variable is outside that range. But if the function ceases to be significant when the variable goes outside a certain range, then the variable is *ipso facto* confined to that range, without the need of any explicit statement to that effect. This principle is to be borne in mind in the development of logical types, to which we shall shortly proceed.

We can now begin to see how it comes that "all so-and-so's" is sometimes a legitimate phrase and sometimes not. Suppose we say "all terms which have the property $\phi$ have the property $\psi$." That means, according to the above interpretation, "$\phi x$ always implies $\psi x$." Provided the range of significance of $\phi x$ is the same as that of $\psi x$, this statement is significant; thus, given any definite function $\phi x$, there are propositions about "all the terms satisfying $\phi x$." But it sometimes happens (as we shall see more fully later on) that what appears verbally as one function is really many analogous functions with different ranges of significance. This applies, for example, to "$p$ is true," which, we shall find, is not really one function of $p$, but is different functions according to the kind of proposition that $p$ is. In such a case, the *phrase* expressing the ambiguous function may, owing to the ambiguity, be significant throughout a set of values of the argument exceeding the range of significance of any one function. In such a case, *all* is not legitimate. Thus if we try to say "all true propositions have the property $\phi$," *i. e.*, "'$p$ is true' always implies $\phi p$," the possible arguments to '$p$ is true' necessarily exceed the possible arguments to $\phi$, and therefore the attempted general statement is impossible. For this reason, genuine general statements about all true propositions can not be made. It may happen, however, that the supposed function $\phi$ is really ambiguous like '$p$ is true;' and if it happens to have an ambiguity precisely of the same kind as that of '$p$ is true,' we may be able always to give an interpretation to the proposition "'$p$ is true' implies $\phi p$." This will occur, *e. g.*, if $\phi p$ is "not-$p$ is false." Thus we get an appearance, in such cases, of a general proposition concerning *all* propositions; but this appearance is due to a systematic ambiguity about such words as *true* and *false*. (This systematic ambiguity results from the hierarchy of propositions which will be explained later on). We may, in all such cases, make our statement about *any* proposition, since the meaning of the ambiguous words will adapt itself to any proposition. But if we turn our proposition into an apparent variable, and say something about *all*, we must suppose the ambiguous words fixed to this or that possible meaning, though it may be quite irrelevant which of their possible meanings they are to have. This is how it happens both that *all* has limitations which exclude "all propositions," and that there nevertheless *seem* to be true statements about "all propositions." Both these points will become plainer when the theory of types has been explained.

It has often been suggested* that what is required in order that it may be legitimate to speak of *all* of a collection is that the collection should be finite. Thus "all men are mortal" will be legitimate because men form a finite class. But that is not really the reason why we can speak of "all men." What is essential, as appears from the above discussion, is not finitude, but what may be called *logical homogeneity.* This property is to belong to any collection whose terms are all contained within the range of significance of some one function. It would always be obvious at a glance whether a collection possessed this property or not, if it were not for the concealed ambiguity in common logical terms such as *true* and *false,* which gives an appearance of being a single function to what is really a conglomeration of many functions with different ranges of significance.

The conclusions of this section are as follows: Every proposition containing *all* asserts that some propositional function is always true; and this means that all values of the said function are true, not that the function is true for all arguments, since there are arguments for which any given function is meaningless, *i. e.*, has no value. Hence we can speak of *all* of a collection when and only when the collection forms part or the whole of the *range of significance* of some propositional function, the range of significance being defined as the collection of those arguments for which the function in question is significant, *i. e.*, has a value.

## IV.

### *The Hierarchy of Types.*

A *type* is defined as the range of significance of a propositional function, *i. e.*, as the collection of arguments for which the said function has values. Whenever an apparent variable occurs in a proposition, the range of values of the apparent variable is a type, the type being fixed by the function of which "all

**E. g.*, by M. Poincaré, *Revue de Métaphysique et de Morale*, Mai, 1906.

values" are concerned. The division of objects into types is necessitated by the reflexive fallacies which otherwise arise. These fallacies, as we saw, are to be avoided by what may be called the "vicious-circle principle;" *i. e.*, "no totality can contain members defined in terms of itself." This principle, in our technical language, becomes: "Whatever contains an apparent variable must not be a possible value of that variable." Thus whatever contains an apparent variable must be of a different type from the possible values of that variable; we will say that it is of a *higher* type. Thus the apparent variables contained in an expression are what determines its type. This is the guiding principle in what follows.

Propositions which contain apparent variables are generated from such as do not contain these apparent variables by processes of which one is always the process of *generalization*, *i. e.*, the substitution of a variable for one of the terms of a proposition, and the assertion of the resulting function for all possible values of the variable. Hence a proposition is called a *generalized* proposition when it contains an apparent variable. A proposition containing no apparent variable we will call an *elementary* proposition. It is plain that a proposition containing an apparent variable presupposes others from which it can be obtained by generalization; hence all generalized propositions presuppose elementary propositions. In an elementary proposition we can distinguish one or more *terms* from one or more *concepts;* the *terms* are whatever can be regarded as the *subject* of the proposition, while the concepts are the predicates or relations asserted of these terms.* The terms of elementary propositions we will call *individuals;* these form the first or lowest type.

It is unnecessary, in practice, to know what objects belong to the lowest type, or even whether the lowest type of variable occurring in a given context is that of individuals or some other. For in practice only the *relative* types of variables are relevant; thus the lowest type occurring in a given context may be called that of individuals, so far as that context is concerned. It follows that the above account of individuals is not essential to the truth of what follows; all that is essential is the way in which other types are generated from individuals, however the type of individuals may be constituted.

By applying the process of generalization to individuals occurring in elementary propositions, we obtain new propositions. The legitimacy of this

*See *Principles of Mathematics*, §48.

process requires only that no individuals should be propositions. That this is so, is to be secured by the meaning we give to the word *individual.* We may define an individual as something destitute of complexity; it is then obviously not a proposition, since propositions are essentially complex. Hence in applying the process of generalization to individuals we run no risk of incurring reflexive fallacies.

Elementary propositions together with such as contain only individuals as apparent variables we will call *first-order propositions.* These form the second logical type.

We have thus a new totality, that of *first-order propositions.* We can thus form new propositions in which first-order propositions occur as apparent variables. These we will call *second-order propositions;* these form the third logical type. Thus, *e. g.*, if Epimenides asserts "all first-order propositions affirmed by me are false," he asserts a second-order proposition; he may assert this truly, without asserting truly any first-order proposition, and thus no contradiction arises.

The above process can be continued indefinitely. The $n+1$th logical type will consist of propositions of order $n$, which will be such as contain propositions of order $n-1$, but of no higher order, as apparent variables. The types so obtained are mutually exclusive, and thus no reflexive fallacies are possible so long as we remember that an apparent variable must always be confined within some one type.

In practice, a hierarchy of *functions* is more convenient than one of propositions. Functions of various orders may be obtained from propositions of various orders by the method of *substitution.* If $p$ is a proposition, and $a$ a constituent of $p$, let "$p/a \,;\, x$" denote the proposition which results from substituting $x$ for $a$ wherever $a$ occurs in $p$. Then $p/a$, which we will call a *matrix,* may take the place of a function; its value for the argument $x$ is $p/a \,;\, x$, and its value for the argument $a$ is $p$. Similarly, if "$p/(a, b) \,;\, (x, y)$" denotes the result of first substituting $x$ for $a$ and then substituting $y$ for $b$, we may use the double matrix $p/(a, b)$ to represent a double function. In this way we can avoid apparent variables other than individuals and propositions of various orders. The *order* of a matrix will be defined as being the order of the proposition in which the substitution is effected, which proposition we will call the *prototype.* The order of a matrix does not determine its type: in the first place because it does not determine the number of arguments for which others are to be substituted (*i. e.*, whether the matrix is of the form $p/a$ or $p/(a, b)$ or $p/(a, b, c)$ etc.); in the second place because, if the prototype is of more than the first order, the arguments may be either propositions or individuals. But it is plain that the type of a matrix is definable always by means of the hierarchy of propositions.

Although it is *possible* to replace functions by matrices, and although this procedure introduces a certain simplicity into the explanation of types, it is technically inconvenient. Technically, it is convenient to replace the prototype $p$ by $\phi a$, and to replace $p/a ; x$ by $\phi x$; thus where, if matrices were being employed, $p$ and $a$ would appear as apparent variables, we now have $\phi$ as our apparent variable. In order that $\phi$ may be legitimate as an apparent variable, it is necessary that its values should be confined to propositions of some one type. Hence we proceed as follows.

A function whose argument is an individual and whose value is always a first-order proposition will be called a first-order function. A function involving a first-order function or proposition as apparent variable will be called a second-order function, and so on. A function of one variable which is of the order next above that of its argument will be called a *predicative* function; the same name will be given to a function of several variables if there is one among these variables in respect of which the function becomes predicative when values are assigned to all the other variables. Then the type of a function is determined by the type of its values and the number and type of its arguments.

The hierarchy of functions may be further explained as follows. A first-order function of an individual $x$ will be denoted by $\phi ! x$ (the letters $\psi$, $\chi$, $\theta$, $f$, $g$, $F$, $G$ will also be used for functions). No first-order function contains a function as apparent variable; hence such functions form a well-defined totality, and the $\phi$ in $\phi ! x$ can be turned into an apparent variable. Any proposition in which $\phi$ appears as apparent variable, and there is no apparent variable of higher type than $\phi$, is a second-order proposition. If such a proposition contains an individual $x$, it is not a predicative function of $x$; but if it contains a first-order function $\phi$, it is a predicative function of $\phi$, and will be written $f ! (\psi ! \hat{z})$. Then $f$ is a *second-order predicative function*; the possible values of $f$ again form a well-defined totality, and we can turn $f$ into an apparent variable. We can thus define *third-order predicative functions*, which will be such as have third-order propositions for their values and second-order predicative functions for their arguments. And in this way we can proceed indefinitely. A precisely similar development applies to functions of several variables.

We will adopt the following conventions. Variables of the lowest type occurring in any context will be denoted by small Latin letters (excluding $f$ and $g$, which are reserved for functions); a predicative function of an argument $x$ (where $x$ may be of any type) will be denoted by $\phi ! x$ (where $\psi$, $\chi$, $\theta$, $f$, $g$, $F$ or $G$ may replace $\phi$); similarly a predicative function of two arguments $x$ and $y$ will be denoted by $\phi ! (x, y)$; a general function of $x$ will be denoted by $\phi x$, and a general function of $x$ and $y$ by $\phi(x, y)$. In $\phi x$, $\phi$ can not be made into an apparent variable, since its type is indeterminate; but in $\phi ! x$, where $\phi$ is a *predicative* function whose argument is of some given type, $\phi$ *can* be made into an apparent variable.

It is important to observe that since there are various types of propositions and functions, and since generalization can only be applied within some one type, all phrases containing the words "all propositions" or "all functions" are *primâ facie* meaningless, though in certain cases they are capable of an unobjectionable interpretation. The contradictions arise from the use of such phrases in cases where no innocent meaning can be found.

If we now revert to the contradictions, we see at once that some of them are solved by the theory of types. Wherever "all propositions" are mentioned, we must substitute "all propositions of order $n$," where it is indifferent what value we give to $n$, but it is essential that $n$ should have *some* value. Thus when a man says "I am lying," we must interpret him as meaning: "There is a proposition of order $n$, which I affirm, and which is false." This is a proposition of order $n + 1$; hence the man is not affirming any proposition of order $n$; hence his statement is false, and yet its falsehood does not imply, as that of "I am lying" appeared to do, that he is making a true statement. This solves the liar.

Consider next "the least integer not nameable in fewer than nineteen syllables." It is to be observed, in the first place, that *nameable* must mean "nameable by means of such-and-such assigned names," and that the number of assigned names must be finite. For if it is not finite, there is no reason why there should be any integer not nameable in fewer than nineteen syllables, and the paradox collapses. We may next suppose that "nameable in terms of names of the class $N$" means "being the only term satisfying some function composed wholly of names of the class $N$." The solution of this paradox lies, I think, in the simple observation that "nameable in terms of names of the class $N$" is never itself nameable in terms of names of that class. If we enlarge $N$ by adding the name "nameable in terms of names of the class $N$," our fundamental apparatus of names is enlarged; calling the new apparatus $N'$, "nameable in terms of names of the class $N'$" remains not nameable in terms of names of the class $N'$. If we try to enlarge $N$ till it embraces *all* names, "nameable" becomes (by what was said above) "being the only term satisfying some function composed wholly of names." But here there is a function as apparent variable; hence we are confined to predicative functions of some one type (for non-predicative functions can not be apparent variables). Hence we have only to observe that nameability in terms of such functions is non-predicative in order to escape the paradox.

The case of "the least indefinable ordinal" is closely analogous to the case we have just discussed. Here, as before, "definable" must be relative to some given apparatus of fundamental ideas; and there is reason to suppose that "definable in terms of ideas of the class $N$" is not definable in terms of ideas of the class $N$. It will be true that there is some definite segment of the series of ordinals consisting wholly of definable ordinals, and having the least indefinable ordinal as its limit. This least indefinable ordinal will be definable by a slight enlargement of our fundamental apparatus; but there will then be a new ordinal which will be the least that is indefinable with the new apparatus. If we enlarge our apparatus so as to include all possible ideas, there is no longer any reason to believe that there is any indefinable ordinal. The apparent force of the paradox lies largely, I think, in the supposition that if all the ordinals of a certain class are definable, the class must be definable, in which case its successor is of course also definable; but there is no reason for accepting this supposition.

The other contradictions, that of Burali-Forti in particular, require some further developments for their solution.

## V.

### *The Axiom of Reducibility.*

A propositional function of $x$ may, as we have seen, be of any order; hence any statement about "all properties of $x$" is meaningless. (A "property of $x$" is the same thing as a "propositional function which holds of $x$.") But it is absolutely necessary, if mathematics is to be possible, that we should have some method of making statements which will usually be equivalent to what we have in mind when we (inaccurately) speak of "all properties of $x$." This necessity appears in many cases, but especially in connection with mathematical induction. We can say, by the use of *any* instead of *all*, "Any property possessed by 0, and by the successors of all numbers possessing it, is possessed by all finite numbers." But we can not go on to: "A finite number is one which possesses *all* properties possessed by 0 and by the successors of all numbers possessing them." If we confine this statement to all first-order properties of numbers, we can not infer that it holds of second-order properties. For example, we shall be unable to prove that if $m$, $n$ are finite numbers, then $m + n$ is a finite number. For, with the above definition, "$m$ is a finite number" is a second-order property of $m$; hence the fact that $m + 0$ is a finite number, and that, if $m + n$ is a finite number, so is $m + n + 1$, does not allow us to conclude by induction that $m + n$ is a finite number. It is obvious that such a state of things renders much of elementary mathematics impossible.

The other definition of finitude, by the non-similarity of whole and part, fares no better. For this definition is: "A class is said to be finite when every one-one relation whose domain is the class and whose converse domain is contained in the class has the whole class for its converse domain." Here a variable relation appears, *i. e.*, a variable function of two variables; we have to take *all* values of this function, which requires that it should be of some assigned order; but any assigned order will not enable us to deduce many of the propositions of elementary mathematics.

Hence we must find, if possible, some method of reducing the order of a propositional function without affecting the truth or falsehood of its values. This seems to be what common-sense effects by the admission of *classes*. Given any propositional function $\phi x$, of whatever order, this is assumed to be equivalent, for all values of $x$, to a statement of the form "$x$ belongs to the class $\alpha$." Now this statement is of the first order, since it makes no allusion to "all functions of such-and-such a type." Indeed its only practical advantage over the original statement $\phi x$ is that it is of the first order. There is no advantage in assuming that there really are such things as classes, and the contradiction about the classes which are not members of themselves shows that, if there are classes, they must be something radically different from individuals. I believe the chief purpose which classes serve, and the chief reason which makes them linguistically convenient, is that they provide a method of reducing the order of a propositional function. I shall, therefore, not assume anything of what may seem to be involved in the common-sense admission of classes, except this: that every propositional function is equivalent, for all its values, to some predicative function.

This assumption with regard to functions is to be made whatever may be the type of their arguments. Let $\phi x$ be a function, of any order, of an argument $x$, which may itself be either an individual or a function of any order. If $\phi$ is of the order next above $x$, we write the function in the form $\phi ! x$; in such a case we will call $\phi$ a *predicative* function. Thus a predicative function of an individual is a first-order function; and for higher types of arguments, predicative functions take the place that first-order functions take in respect of individuals. We assume, then, that every function is equivalent, for all its values, to some predicative function of the same argument. This assumption seems to be the essence of the usual assumption of classes; at any rate, it retains as much of classes as we have any use for, and little enough to avoid the contradictions which a less grudging admission of classes is apt to entail. We will call this assumption the *axiom of classes*, or the *axiom of reducibility.*

We shall assume similarly that every function of two variables is equivalent, for all its values, to a predicative function of those variables, where a predicative function of two variables is one such that there is one of the variables in respect of which the function becomes predicative (in our previous sense) when a value is assigned to the other variable. This assumption is what seems to be meant by saying that any statement about two variables defines a relation between them. We will call this assumption the *axiom of relations* or the *axiom of reducibility.*

In dealing with relations between more than two terms, similar assumptions would be needed for three, four, ... variables. But these assumptions are not indispensable for our purpose, and are therefore not made in this paper.

By the help of the axiom of reducibility, statements about "all first-order functions of $x$" or "all predicative functions of $\alpha$" yield most of the results which otherwise would require "all functions." The essential point is that such results are obtained in all cases where only the truth or falsehood of values of the functions concerned are relevant, as is invariably the case in mathematics. Thus mathematical induction, for example, need now only be stated for all predicative functions of numbers; it then follows from the axiom of classes that it holds of *any* function of whatever order. It might be thought that the paradoxes for the sake of which we invented the hierarchy of types would now reappear. But this is not the case, because, in such paradoxes, either something beyond the truth or falsehood of values of functions is relevant, or expressions occur which are unmeaning even after the introduction of the axiom of reducibility. For example, such a statement as "Epimenides asserts $\psi x$" is not equivalent to "Epimenides asserts $\phi ! x$," even though $\psi x$ and $\phi ! x$ are equivalent. Thus "I am lying" remains unmeaning if we attempt to include *all* propositions among those which I may be falsely affirming, and is unaffected by the axiom of classes if we confine it to propositions of order $n$. The hierarchy of propositions and functions, therefore, remains relevant in just those cases in which there is a paradox to be avoided.

## VI.

### *Primitive Ideas and Propositions of Symbolic Logic.*

The primitive ideas required in symbolic logic appear to be the following seven:

(1) Any propositional function of a variable $x$ or of several variables $x, y, z, \ldots$ This will be denoted by $\phi x$ or $\phi(x, y, z, \ldots)$

(2) The negation of a proposition. If $p$ is the proposition, its negation will be denoted by $\sim p$.

(3) The disjunction or logical sum of two propositions; *i. e.*, "this or that." If $p$, $q$ are the two propositions, their disjunction will be denoted by $p \vee q$.*

(4) The truth of *any* value of a propositional function; *i. e.*, of $\phi x$ where $x$ is not specified.

(5) The truth of *all* values of a propositional function. This is denoted by $(x) . \phi x$ or $(x) : \phi x$ or whatever larger number of dots may be necessary to bracket off the proposition.† In $(x) . \phi x$, $x$ is called an *apparent variable*, whereas when $\phi x$ is asserted, where $x$ is not specified, $x$ is called a *real variable.*

(6) Any predicative function of an argument of any type; this will be represented by $\phi ! x$ or $\phi ! \alpha$ or $\phi ! R$, according to circumstances. A predicative function of $x$ is one whose values are propositions of the type next above that of $x$, if $x$ is an individual or a proposition, or that of values of $x$ if $x$ is a

---

*In a previous article in this journal, I took implication as indefinable, instead of disjunction. The choice between the two is a matter of taste; I now choose disjunction, because it enables us to diminish the number of primitive propositions.

†The use of dots follows Peano's usage. It is fully explained by Mr. Whitehead, "On Cardinal Numbers," AMERICAN JOURNAL OF MATHEMATICS, Vol. XXIV, and "On Mathematical Concepts of the Material World," *Phil. Trans. A.*, Vol. CCV, p. 472.

function. It may be described as one in which the apparent variables, if any, are all of the same type as $x$ or of lower type; and a variable is of lower type than $x$ if it can significantly occur as argument to $x$, or as argument to an argument to $x$, etc.

(7) Assertion; *i. e.*, the assertion that some proposition is true, or that any value of some propositional function is true. This is required to distinguish a proposition actually asserted from one merely considered, or from one adduced as hypothesis to some other. It will be indicated by the sign "$\vdash$" prefixed to what is asserted, with enough dots to bracket off what is asserted.*

Before proceeding to the primitive propositions, we need certain definitions. In the following definitions, as well as in the primitive propositions, the letters $p$, $q$, $r$ are used to denote propositions.

$$p \supset q \,.=.\, \sim p \vee q \quad \text{Df.}$$

This definition states that "$p \supset q$" (which is read "$p$ implies $q$") is to mean "$p$ is false or $q$ is true." I do not mean to affirm that "implies" can not have any other meaning, but only that this meaning is the one which it is most convenient to give to "implies" in symbolic logic. In a definition, the sign of equality and the letters "Df" are to be regarded as one symbol, meaning jointly "is defined to mean." The sign of equality without the letters "Df" has a different meaning, to be defined shortly.

$$p \,.\, q \,.=.\, \sim(\sim p \vee \sim q) \quad \text{Df.}$$

This defines the logical product of two propositions $p$ and $q$, *i. e.*, "$p$ and $q$ are both true." The above definition states that this is to mean: "It is false that either $p$ is false or $q$ is false." Here again, the definition does not give the only meaning which can be given to "$p$ and $q$ are both true," but gives the meaning which is most convenient for our purposes.

$$p \equiv q \,.=.\, p \supset q \,.\, q \supset p \quad \text{Df.}$$

That is, "$p \equiv q$," which is read "$p$ is equivalent to $q$," means "$p$ implies $q$ and $q$ implies $p$;" whence, of course, it follows that $p$ and $q$ are both true or both false.

$$(\exists x) \,.\, \phi x \,.=.\, \sim\{(x) \,.\, \sim \phi x\} \quad \text{Df.}$$

---

* This sign, as well as the introduction of the idea which it expresses, are due to Frege. See his *Begriffsschrift* (Halle, 1879), p. 1, and *Grundgesetze der Arithmetik*, Vol. I (Jena, 1893), p. 9.

This defines "there is at least one value of $x$ for which $\phi x$ is true." We define it as meaning "it is false that $\phi x$ is always false."

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

This is the definition of identity. It states that $x$ and $y$ are to be called identical when every predicative function satisfied by $x$ is satisfied by $y$. It follows from the axiom of reducibility that if $x$ satisfies $\psi x$, where $\psi$ is any function, predicative or non-predicative, then $y$ satisfies $\psi y$.

The following definitions are less important, and are introduced solely for the purpose of abbreviation.

$$\begin{array}{lll}
(x, y) \,.\, \phi(x, y) \,. & = : (x) : (y) \,.\, \phi(x, y) & \text{Df.} \\
(\exists x, y) \,.\, \phi(x, y) \,. & = : (\exists x) : (\exists y) \,.\, \phi(x, y) & \text{Df.} \\
\phi x \,.\, \supset_x \,.\, \psi x : & = : (x) : \phi x \supset \psi x & \text{Df.} \\
\phi x \,.\, \equiv_x \,.\, \psi x : & = : (x) : \phi x \,.\, \equiv \,.\, \psi x & \text{Df.} \\
\phi(x, y) \,.\, \supset_{x, y} \,.\, \psi(x, y) : & = : (x, y) : \phi(x, y) \,.\, \supset \,.\, \psi(x, y) & \text{Df.,}
\end{array}$$

and so on for any number of variables.

The primitive propositions required are as follows. (In 2, 3, 4, 5, 6, and 10, $p$, $q$, $r$ stand for propositions.)

(1) A proposition implied by a true premise is true.

(2) $\vdash : p \vee p \,.\, \supset . \, p$.

(3) $\vdash : q \,.\, \supset . \, p \vee q$.

(4) $\vdash : p \vee q \,.\, \supset . \, q \vee p$.

(5) $\vdash : p \vee (q \vee r) \,.\, \supset . \, q \vee (p \vee r)$.

(6) $\vdash :. \, q \supset r \,.\, \supset : p \vee q \,.\, \supset . \, p \vee r$.

(7) $\vdash : (x) \,.\, \phi x \,.\, \supset . \, \phi y$;

*i. e.*, "if all values of $\phi \hat{x}$ are true, then $\phi y$ is true, where $\phi y$ is any value." *

(8) If $\phi y$ is true, where $\phi y$ is any value of $\phi \hat{x}$, then $(x) \,.\, \phi x$ is true. This can not be expressed in our symbols; for if we write "$\phi y \,.\, \supset . \,(x) \,.\, \phi x$," that means "$\phi y$ implies that all values of $\phi \hat{x}$ are true, where $y$ may have any value of the appropriate type," which is not in general the case. What we mean to assert is: "If, however $y$ is chosen, $\phi y$ is true, then $(x) \,.\, \phi x$ is true;" whereas what is expressed by "$\phi y \,.\, \supset . \,(x) \,.\, \phi x$" is: "However $y$ is chosen, if $\phi y$ is true, then $(x) \,.\, \phi x$ is true," which is quite a different statement, and in general a false one.

* It is convenient to use the notation $\phi \hat{x}$ to denote the function itself, as opposed to this or that value of the function.

(9) $\vdash : (x) . \phi x . \supset . \phi a$, where $a$ is any definite constant.

This principle is really as many different principles as there are possible values of $a$. *I. e.*, it states that, *e. g.*, whatever holds of all individuals holds of Socrates; also that it holds of Plato; and so on. It is the principle that a general rule may be applied to particular cases; but in order to give it scope, it is necessary to mention the particular cases, since otherwise we need the principle itself to assure us that the general rule that general rules may be applied to particular cases may be applied (say) to the particular case of Socrates. It is thus that this principle differs from (7); our present principle makes a statement about Socrates, or about Plato, or some other definite constant, whereas (7) made a statement about a variable.

The above principle is never used in symbolic logic or in pure mathematics, since all our propositions are general, and even when (as in "one is a number") we seem to have a strictly particular case, this turns out not to be so when closely examined. In fact, the use of the above principle is the distinguishing mark of *applied* mathematics. Thus, strictly speaking, we might have omitted it from our list.

(10) $\vdash :. (x) . p \vee \phi x . \supset : p . \vee . (x) . \phi x$;

*i. e.*, "if '$p$ or $\phi x$' is always true, then either $p$ is true, or $\phi x$ is always true."

(11) When $f(\phi x)$ is true whatever argument $x$ may be, and $F(\phi y)$ is true whatever possible argument $y$ may be, then $\{f(\phi x) . F(\phi x)\}$ is true whatever possible argument $x$ may be.

This is the axiom of the "identification of variables." It is needed when two separate propositional functions are each known to be always true, and we wish to infer that their logical product is always true. This inference is only legitimate if the two functions take arguments of the same type, for otherwise their logical product is meaningless. In the above axiom, $x$ and $y$ must be of the same type, because both occur as arguments to $\phi$.

(12) If $\phi x . \phi x \supset \psi x$ is true for any possible $x$, then $\psi x$ is true for any possible $x$.

This axiom is required in order to assure us that the range of significance of $\psi x$, in the case supposed, is the same as that of $\phi x . \phi x \supset \psi x . \supset . \psi x$; both are in fact the same as that of $\phi x$. We know, in the case supposed, that $\psi x$ is true whenever $\phi x . \phi x \supset \psi x$ and $\phi x . \phi x \supset \psi x . \supset . \psi x$ are both significant, but we do not know, without an axiom, that $\psi x$ is true whenever $\psi x$ is significant. Hence the need of the axiom.

Axioms (11) and (12) are required, *e. g.*, in proving

$$(x) . \phi x : (x) . \phi x \supset \psi x : \supset . (x) . \psi x.$$

By (7) and (11),

$$\vdash :. (x) . \phi x : (x) . \phi x \supset \psi x : \supset : \phi y . \phi y \supset \psi y$$

whence by (12),

$$\vdash :. (x) . \phi x : (x) . \phi x \supset \psi x : \supset : \psi y,$$

whence the result follows by (8) and (10).

(13) $\vdash :. (\exists f) :. (x) : \phi x . \equiv . f ! x.$

This is the axiom of reducibility. It states that, given any function $\phi \hat{x}$, there is a predicative function $f ! \hat{x}$ such that $f ! x$ is always equivalent to $\phi x$. Note that, since a proposition beginning with "$(\exists f)$" is, by definition, the negation of one beginning with "$(f)$," the above axiom involves the possibility of considering "all predicative functions of $x$." If $\phi x$ is *any* function of $x$, we can not make propositions beginning with "$(\phi)$" or "$(\exists \phi)$," since we can not consider "all functions," but only "*any* function" or "all *predicative* functions."

(14) $\vdash :. (\exists f) :. (x, y) : \phi(x, y) . \equiv . f ! (x, y).$

This is the axiom of reducibility for double functions.

In the above propositions, our $x$ and $y$ may be of any type whatever. The only way in which the theory of types is relevant is that (11) only allows us to identify real variables occurring in different contents when they are shown to be of the same type by both occurring as arguments to the same function, and that, in (7) and (9), $y$ and $a$ must respectively be of the appropriate type for arguments to $\phi \hat{z}$. Thus, for example, suppose we have a proposition of the form $(\phi) . f ! (\phi ! \hat{z}, x)$, which is a second-order function of $x$. Then by (7),

$$\vdash : (\phi) . f ! (\phi ! \hat{z}, x) . \supset . f ! (\psi ! \hat{z}, x),$$

where $\psi ! \hat{z}$ is any *first*-order function. But it will not do to treat $(\phi) . f ! (\phi ! \hat{z}, x)$ as if it were a first-order function of $x$, and take this function as a possible value of $\psi ! \hat{z}$ in the above. It is such confusions of types that give rise to the paradox of the *liar*.

Again, consider the classes which are not members of themselves. It is plain that, since we have identified classes with functions,* no class can be significantly said to be or not to be a member of itself; for the members of a class are arguments to it, and arguments to a function are always of lower type

* This identification is subject to a modification to be explained shortly.

than the function. And if we ask: "But how about the class of all classes? Is not that a class, and so a member of itself?", the answer is twofold. First, if "the class of all classes" means "the class of all classes of whatever type," then there is no such notion. Secondly, if "the class of all classes" means "the class of all classes of type $t$," then this is a class of the next type above $t$, and is therefore again not a member of itself.

Thus although the above primitive propositions apply equally to all types, they do not enable us to elicit contradictions. Hence in the course of any deduction it is never necessary to consider the absolute type of a variable; it is only necessary to see that the different variables occurring in one proposition are of the proper relative types. This excludes such functions as that from which our fourth contradiction was obtained, namely: "The relation $R$ holds between $R$ and $S$." For a relation between $R$ and $S$ is necessarily of higher type than either of them, so that the proposed function is meaningless.

## VII.

### *Elementary Theory of Classes and Relations.*

Propositions in which a function $\phi$ occurs may depend, for their truth-value, upon the particular function $\phi$, or they may depend only upon the *extension* of $\phi$, *i. e.*, upon the arguments which satisfy $\phi$. A function of the latter sort we will call *extensional.* Thus, *e. g.*, "I believe that all men are mortal" may not be equivalent to "I believe that all featherless bipeds are mortal," even if men are coextensive with featherless bipeds; for I may not know that they are coextensive. But "all men are mortal" must be equivalent to "all featherless bipeds are mortal" if men are coextensive with featherless bipeds. Thus "all men are mortal" is an extensional function of the function "$x$ is a man," while "I believe all men are mortal" is a function which is not extensional; we will call functions *intensional* when they are not extensional. The functions of functions with which mathematics is specially concerned are all extensional. The mark of an extensional function $f$ of a function $\phi\,!\,\hat{z}$ is

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

From any function $f$ of a function $\phi\,!\,\hat{z}$ we can derive an associated extensional function as follows. Put

$$f\{\hat{z}(\psi z)\} \,.=:\, (\exists\phi) : \phi\,!\,x \,.\equiv_x.\, \psi x : f\{\phi\,!\,\hat{z}\} \quad \text{Df.}$$

The function $f\{\hat{z}(\psi z)\}$ is in reality a function of $\psi\hat{z}$, though not the same function as $f(\psi\hat{z})$, supposing this latter to be significant. But it is convenient to treat $f\{\hat{z}(\psi z)\}$ technically as though it had an argument $\hat{z}(\psi z)$, which we call "the class defined by $\psi$." We have

$$\vdash :. \phi x . \equiv_x . \psi x : \supset : f\{\hat{z}(\phi z)\} . \equiv . f\{\hat{z}(\psi z)\},$$

whence, applying to the fictitious objects $\hat{z}(\phi z)$ and $\hat{z}(\psi z)$ the definition of identity given above, we find

$$\vdash :. \phi x . \equiv_x . \psi x : \supset . \hat{z}(\phi z) = \hat{z}(\psi z).$$

This, with its converse (which can also be proved), is the distinctive property of classes. Hence we are justified in treating $\hat{z}(\phi z)$ as the class defined by $\phi$. In the same way we put

$$f\{\hat{x}\hat{y}\,\psi(x, y)\} . = : (\exists\phi) : \phi\,!\,(x, y) . \equiv_{x,y} . \psi(x, y) : f\{\phi\,!\,(\hat{x}, \hat{y})\} \quad \text{Df.}$$

A few words are necessary here as to the distinction between $\phi\,!\,(\hat{x}, \hat{y})$ and $\phi\,!\,(\hat{y}, \hat{x})$. We will adopt the following convention: When a function (as opposed to its values) is represented in a form involving $\hat{x}$ and $\hat{y}$, or any other two letters of the alphabet, the value of this function for the arguments $a$ and $b$ is to be found by substituting $a$ for $\hat{x}$ and $b$ for $\hat{y}$; *i. e.*, the argument mentioned first is to be substituted for the letter which comes earlier in the alphabet, and the argument mentioned second for the later letter. This sufficiently distinguishes between $\phi\,!\,(\hat{x}, \hat{y})$ and $\phi\,!\,(\hat{y}, \hat{x})$; *e. g.*:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| The | value of | $\phi\,!\,(\hat{x}, \hat{y})$ | for | arguments | $a, b$ | is | $\phi\,!\,(a, b)$. |
| " | " | " | " | " | $b, a$ | " | $\phi\,!\,(b, a)$. |
| " | " | $\phi\,!\,(\hat{y}, \hat{x})$ | " | " | $a, b$ | " | $\phi\,!\,(b, a)$. |
| " | " | " | " | " | $b, a$ | " | $\phi\,!\,(a, b)$. |

We put

$$x \epsilon \phi\,!\,\hat{z} . = . \phi\,!\,x \quad \text{Df.},$$

whence

$$\vdash :. x \epsilon \hat{z}(\psi z) . = : (\exists\phi) : \phi\,!\,y . \equiv_y . \psi y : \phi\,!\,x.$$

Also by the reducibility-axiom we have

$$(\exists\phi) : \phi\,!\,y . \equiv_y . \psi y,$$

whence

$$\vdash : x \epsilon \hat{z}(\psi z) . \equiv . \psi x.$$

This holds whatever $x$ may be. Suppose now we want to consider $\hat{z}(\psi z)\,\varepsilon\hat{\phi}f\{\hat{z}(\phi\,!\,z)\}$. We have, by the above,

$$\vdash :.\ \hat{z}(\psi z)\,\varepsilon\hat{\phi}f\{\hat{z}(\phi\,!\,z)\}\ .\equiv: f\{\hat{z}(\psi z)\} :\equiv: (\exists\phi) : \phi\,!\,y\ .\equiv_y .\ \psi y : f\{\phi\,!\,z\},$$

whence

$$\vdash :.\ \hat{z}(\psi z)=\hat{z}(\chi z)\ .\supset: \hat{z}(\psi z)\varepsilon\kappa\ .\equiv_\kappa .\ \hat{z}(\chi z)\varepsilon\kappa,$$

where $\kappa$ is written for any expression of the form $\hat{\phi}f\{\hat{z}(\phi\,!\,z)\}$.

We put

$$cls = \hat{\alpha}\{(\exists\phi)\,.\,\alpha=\hat{z}(\phi\,!\,z)\}\quad \text{Df.}$$

Here $cls$ has a meaning which depends upon the type of the apparent variable $\phi$. Thus, *e. g.*, the proposition "$cls\ \varepsilon\ cls$," which is a consequence of the above definition, requires that "$cls$" should have a different meaning in the two places where it occurs. The symbol "$cls$" can only be used where it is unnecessary to know the type; it has an ambiguity which adjusts itself to circumstances. If we introduce as an indefinable the function "Indiv ! $x$," meaning "$x$ is an individual," we may put

$$Kl = \hat{\alpha}\{(\exists\phi)\,.\,\alpha=\hat{z}(\phi\,!\,z\,.\,\mathrm{Indiv}\,!\,z)\}\quad \text{Df.}$$

Then $Kl$ is an unambiguous symbol meaning "classes of individuals."

We will use small Greek letters (other than $\varepsilon$, $\phi$, $\psi$, $\chi$, $\theta$) to represent classes of whatever type; *i. e.*, to stand for symbols of the form $\hat{z}(\phi\,!\,z)$ or $\hat{z}(\phi z)$.

The theory of classes proceeds, from this point on, much as in Peano's system; $\hat{z}(\phi z)$ replaces $z\ni(\phi z)$. Also I put

$$\alpha \subset \beta\ .=: x\varepsilon\alpha\ .\supset_x .\ x\varepsilon\beta \quad \text{Df.}$$

$$\exists\,!\,\alpha\ .=.\ (\exists x)\,.\,x\varepsilon\alpha \quad \text{Df.}$$

$$V=\hat{x}(x=x) \quad \text{Df.}$$

$$\Lambda = x\{\sim(x=x)\} \quad \text{Df.}$$

where $\Lambda$, as with Peano, is the null-class. The symbols $\exists$, $\Lambda$, $V$, like $cls$ and $\varepsilon$, are ambiguous, and only acquire a definite meaning when the type concerned is otherwise indicated.

We treat relations in exactly the same way, putting

$$a\{\phi\,!\,(\hat{x},\hat{y})\}b\ .=.\ \phi\,!\,(a,b) \quad \text{Df.}$$

(the order being determined by the alphabetical order of $x$ and $y$ and the typographical order of $a$ and $b$); whence

$$\vdash :. a\{\hat{x}\hat{y}\psi(x, y)\}b . \equiv : (\exists\phi) : \psi(x, y) . \equiv_{x, y} . \phi!(x, y) : \phi!(a, b),$$

whence, by the reducibility-axiom,

$$\vdash : a\{\hat{x}\hat{y}\psi(x, y)\}b . \equiv . \psi(a, b).$$

We use Latin capital letters as abbreviations for such symbols as $\hat{x}\hat{y}\psi(x, y)$, and we find

$$\vdash :. R = S . \equiv : xRy . \equiv_{x, y} . xSy,$$

where

$$R = S . = : f!R . \supset_f . f!S \quad \text{Df.}$$

We put

$$\text{Rel} = \hat{R}\{(\exists\phi) . R = \hat{x}\hat{y}\phi!(x, y)\} \quad \text{Df.},$$

and we find that everything proved for classes has its analogue for dual relations. Following Peano, we put

$$\alpha \cap \beta = \hat{x}(x\epsilon\alpha . x\epsilon\beta) \quad \text{Df.},$$

defining the product, or common part, of two classes;

$$\alpha \cup \beta = \hat{x}(x\epsilon\alpha . \vee . x\epsilon\beta) \quad \text{Df.},$$

defining the sum of two classes; and

$$-\alpha = \hat{x}\{\sim(x\epsilon\alpha)\} \quad \text{Df.},$$

defining the negation of a class. Similarly for relations we put

$$R \mathbin{\dot{\cap}} S = \hat{x}\hat{y}(xRy . xSy) \quad \text{Df.}$$
$$R \mathbin{\dot{\cup}} S = \hat{x}\hat{y}(xRy . \vee . xSy) \quad \text{Df.}$$
$$\dot{-} R = \hat{x}\hat{y}\{\sim(xRy)\} \quad \text{Df.}$$

## VIII.

### *Descriptive Functions.*

The functions hitherto considered have been propositional functions, with the exception of a few particular functions such $R \mathbin{\dot{\cap}} S$. But the ordinary functions of mathematics, such as $x^2$, $\sin x$, $\log x$, are not propositional. Functions of this kind always mean "the term having such-and-such a relation to $x$." For this reason they may be called *descriptive* functions, because they *describe* a certain term by means of its relation to their argument. Thus "$\sin \pi/2$"

describes the number **1**; yet propositions in which sin $\pi/2$ occurs are not the same as they would be if 1 were substituted. This appears, *e. g.*, from the proposition "sin $\pi/2 = 1$," which conveys valuable information, whereas "$1 = 1$" is trivial. Descriptive functions have no meaning by themselves, but only as constituents of propositions; and this applies generally to phrases of the form "the term having such-and-such a property." Hence in dealing with such phrases, we must define any proposition in which they occur, not the phrases themselves.* We are thus led to the following definition, in which "$(\imath x)\,(\phi x)$" is to be read "*the* term $x$ which satisfies $\phi x$."

$$\psi\{(\imath x)\,(\phi x)\} \,.=:\, (\exists b) : \phi x \,.\equiv_x .\, x = b : \psi b \quad \text{Df.}$$

This definition states that "the term which satisfies $\phi$ satisfies $\psi$" is to mean: "There is a term $b$ such that $\phi x$ is true when and only when $x$ is $b$, and $\psi b$ is true." Thus all propositions about "*the* so-and-so" will be false if there are no so-and-so's or several so-and-so's.

The general definition of a descriptive function is

$$R\text{‘}y = (\imath x)\,(xRy) \quad \text{Df.};$$

that is, "$R\text{‘}y$" is to mean "the term which has the relation $R$ to $y$." If there are several terms or none having the relation $R$ to $y$, all propositions about $R\text{‘}y$ will be false. We put

$$E!\,(\imath x)\,(\phi x) \,.=:\, (\exists b) : \phi x \,.\equiv_x .\, x = b \quad \text{Df.}$$

Here "$E!\,(\imath x)\,(\phi x)$" may be read "there is such a term as the $x$ which satisfies $\phi x$," or "the $x$ which satisfies $\phi x$ exists." We have

$$\vdash :.\, E!\,R\text{‘}y \,.\equiv: (\exists b) : xRy \,.\equiv_x .\, x = b.$$

The inverted comma in $R\text{‘}y$ may be read *of.* Thus if $R$ is the relation of father to son, "$R\text{‘}y$" is "the father of $y$." If $R$ is the relation of son to father, all propositions about $R\text{‘}y$ will be false unless $y$ has one son and no more.

From the above it appears that descriptive functions are obtained from relations. The relations now to be defined are chiefly important on account of the descriptive functions to which they give rise.

$$\text{Cnv} = \hat{Q}\hat{P}\,\{xQy \,.\equiv_{x,y} .\, yPx\} \quad \text{Df.}$$

---

* See the above-mentioned article "On Denoting," where the reasons for this view are given at length.

Here *Cnv* is short for "converse." It is the relation of a relation to its converse; *e. g.*, of *greater* to *less*, of parentage to sonship, of preceding to following, etc. We have

$$\vdash . \mathrm{Cnv}`P = (\imath Q)\{xQy . \equiv_{x,y} . yPx\}.$$

For a shorter notation, often more convenient, we put

$$\breve{P} = \mathrm{Cnv}`P \quad \mathrm{Df.}$$

We want next a notation for the class of terms which have the relation $R$ to $y$. For this purpose, we put

$$\overrightarrow{R} = \hat{\alpha}\hat{y}\{\alpha = \hat{x}(xRy)\} \quad \mathrm{Df.},$$

whence

$$\vdash . \overrightarrow{R}`y = \hat{x}(xRy).$$

Similarly we put

$$\overleftarrow{R} = \hat{\beta}\hat{x}\{\beta = \hat{y}(xRy)\} \quad \mathrm{Df.},$$

whence

$$\vdash . \overleftarrow{R}`x = \hat{y}(xRy).$$

We want next the *domain* of $R$ (*i. e.*, the class of terms which have the relation $R$ to something), the *converse domain* of $R$ (*i. e.*, the class of terms to which something has the relation $R$), and the *field* of $R$, which is the sum of the domain and the converse domain. For this purpose we define the relations of the domain, converse domain, and field, to $R$. The definitions are:

$$D = \hat{\alpha}\hat{R}\{\alpha = \hat{x}((\exists y) . xRy)\} \quad \mathrm{Df.}$$
$$\mathit{C\!\!I} = \hat{\beta}\hat{R}\{\beta = \hat{y}((\exists x) . xRy)\} \quad \mathrm{Df.}$$
$$C = \hat{\gamma}\hat{R}\{\gamma = \hat{x}((\exists y) : xRy . \vee . yRx)\} \quad \mathrm{Df.}$$

Note that the third of these definitions is only significant when $R$ is what we may call a *homogeneous* relation; *i. e.*, one in which, if $xRy$ holds, $x$ and $y$ are of the same type. For otherwise, however we may choose $x$ and $y$, either $xRy$ or $yRx$ will be meaningless. This observation is important in connection with Burali-Forti's contradiction.

We have, in virtue of the above definitions,

$$\vdash . D`R = \hat{x}\{(\exists y) . xRy\},$$
$$\vdash . \mathit{C\!\!I}`R = \hat{y}\{(\exists x) . xRy\},$$
$$\vdash . C`R = \hat{x}\{(\exists y) : xRy . \vee . yRx\},$$

the last of these being significant only when $R$ is homogeneous. "$D$'$R$" is read "the domain of $R$;" "$\text{Ɑ}$'$R$" is read "the converse domain of $R$," and "$C$'$R$" is read "the field of $R$." The letter $C$ is chosen as the initial of the word "campus."

We want next a notation for the relation, to a class $\alpha$ contained in the domain of $R$, of the class of terms to which some member of $\alpha$ has the relation $R$, and also for the relation, to a class $\beta$ contained in the converse domain of $R$, of the class of terms which have the relation $R$ to some member of $\beta$. For the second of these we put

$$R_\epsilon = \hat{\alpha}\hat{\beta}\,\{\alpha = \hat{x}\,((\exists y)\,.\,y\epsilon\beta\,.\,xRy)\} \quad \text{Df.}$$

So that

$$\vdash .\, R_\epsilon\text{'}\beta = \hat{x}\,\{(\exists y)\,.\,y\epsilon\beta\,.\,xRy\}.$$

Thus if $R$ is the relation of father to son, and $\beta$ is the class of Etonians, $R_\epsilon$'$\beta$ will be the class "fathers of Etonians;" if $R$ is the relation "less than," and $\beta$ is the class of proper fractions of the form $1 - 2^{-n}$ for integral values of $n$, $R_\epsilon$'$\beta$ will be the class of fractions less than some fraction of the form $1 - 2^{-n}$; *i. e.*, $R_\epsilon$'$\beta$ will be the class of proper fractions. The other relation mentioned above is $(\breve{R})_\epsilon$.

We put, as an alternative notation often more convenient,

$$R\text{''}\beta = R_\epsilon\text{'}\beta \quad \text{Df.}$$

The *relative product* of two relations $R$, $S$ is the relation which holds between $x$ and $z$ whenever there is a term $y$ such that $xRy$ and $yRz$ both hold. The relative product is denoted by $R \mid S$. Thus

$$R \mid S = \hat{x}\hat{z}\,\{(\exists y)\,.\,xRy\,.\,yRz\} \quad \text{Df.}$$

We put also

$$R^2 = R \mid R \quad \text{Df.}$$

The product and sum of a class of classes are often required. They are defined as follows:

$$s\text{'}\kappa = \hat{x}\,\{(\exists\alpha)\,.\,\alpha\epsilon\kappa\,.\,x\epsilon\alpha\} \quad \text{Df.}$$
$$p\text{'}\kappa = \hat{x}\,\{\alpha\epsilon\kappa\,.\supset_\alpha.\,x\epsilon\alpha\} \quad \text{Df.}$$

Similarly for relations we put

$$\dot{s}\text{'}\lambda = \hat{x}\hat{y}\,\{(\exists R)\,.\,R\epsilon\lambda\,.\,xRy\} \quad \text{Df.}$$
$$\dot{p}\text{'}\lambda = \hat{x}\hat{y}\,\{R\epsilon\lambda\,.\supset_R.\,xRy\} \quad \text{Df.}$$

We need a notation for the class whose only member is $x$. Peano uses $\iota x$, hence we shall use $\iota \text{‘} x$. Peano showed (what Frege also had emphasized) that this class can not be identified with $x$. With the usual view of classes, the need for such a distinction remains a mystery; but with the view set forth above, it becomes obvious.

We put

$$\iota = \hat{\alpha}\hat{x} \{\alpha = \hat{y}\,(y = x)\} \quad \text{Df.},$$

whence

$$\vdash . \iota \text{‘} x = \hat{y}\,(y = x),$$

and

$$\vdash : E ! \breve{\iota} \text{‘} \alpha . \supset . \breve{\iota} \text{‘} \alpha = (\imath x)\,(x \varepsilon \alpha);$$

*i. e.*, if $\alpha$ is a class which has only one member, then $\breve{\iota} \text{‘} \alpha$ is that one member.*

For the class of classes contained in a given class, we put

$$\text{Cl} \text{‘} \alpha = \hat{\beta}\,(\beta \subset \alpha) \quad \text{Df.}$$

We can now proceed to the consideration of cardinal and ordinal numbers, and of how they are affected by the doctrine of types.

## IX.

### *Cardinal Numbers.*

The cardinal number of a class $\alpha$ is defined as the class of all classes *similar* to $\alpha$, two classes being similar when there is a one-one relation between them. The class of one-one relations is denoted by $1 \rightarrow 1$, and defined as follows:

$$1 \rightarrow 1 = \hat{R} \{xRy . x'Ry . xRy' . \supset_{x, y, x', y'} . x = x' . y = y'\} \quad \text{Df.}$$

Similarity is denoted by *Sim*; its definition is

$$\text{Sim} = \hat{\alpha}\hat{\beta} \{(\exists R) . R \varepsilon 1 \rightarrow 1 . D \text{‘} R = \alpha . \text{Œ} \text{‘} R = \beta\} \quad \text{Df.}$$

Then $\overrightarrow{\text{Sim}} \text{‘} \alpha$ is, by definition, the cardinal number of $\alpha$; this we will denote by $Nc \text{‘} \alpha$; hence we put

$$Nc = \overrightarrow{\text{Sim}} \quad \text{Df.},$$

whence

$$\vdash . Nc \text{‘} \alpha = \overrightarrow{\text{Sim}} \text{‘} \alpha.$$

---

*Thus $\breve{\iota} \text{‘} \alpha$ is what Peano calls $\imath a$.

The class of cardinals we will denote by $NC$; thus

$$NC = Nc``cls \quad \text{Df.}$$

0 is defined as the class whose only member is the null-class, $\Lambda$, so that

$$0 = \iota`\Lambda \quad \text{Df.}$$

The definition of 1 is

$$1 = \hat{\alpha}\,\{(\exists c) : x \varepsilon \alpha \,.\equiv_x .\, x = c\} \quad \text{Df.}$$

It is easy to prove that 0 and 1 are cardinals according to the definition.

It is to be observed, however, that 0 and 1 and all the other cardinals, according to the above definitions, are ambiguous symbols, like *cls*, and have as many meanings as there are types. To begin with 0: the meaning of 0 depends upon that of $\Lambda$, and the meaning of $\Lambda$ is different according to the type of which it is the null-class. Thus there are as many 0's as there are types; and the same applies to all the other cardinals. Nevertheless, if two classes $\alpha$, $\beta$ are of different types, we can speak of them as having the same cardinal, or of one as having a greater cardinal than the other, because a one-one relation may hold between the members of $\alpha$ and the members of $\beta$, even when $\alpha$ and $\beta$ are of different types. For example, let $\beta$ be $\iota``\alpha$; *i. e.*, the class whose members are the classes consisting of single members of $\alpha$. Then $\iota``\alpha$ is of higher type than $\alpha$, but similar to $\alpha$, being correlated with $\alpha$ by the one-one relation $\iota$.

The hierarchy of types has important results in regard to addition. Suppose we have a class of $\alpha$ terms and a class of $\beta$ terms, where $\alpha$ and $\beta$ are cardinals; it may be quite impossible to add them together to get a class of $\alpha$ and $\beta$ terms, since, if the classes are not of the same type, their logical sum is meaningless. Where only a finite number of classes are concerned, we can obviate the practical consequences of this, owing to the fact that we can always apply operations to a class which raise its type to any required extent without altering its cardinal number. For example, given any class $\alpha$, the class $\iota``\alpha$ has the same cardinal number, but is of the next type above $\alpha$. Hence, given any finite number of classes of different types, we can raise all of them to the type which is what we may call the lowest common multiple of all the types in question; and it can be shown that this can be done in such a way that the resulting classes shall have no common members. We may then form the logical sum of all the classes so obtained, and its cardinal number will be the arithmetical sum of the cardinal numbers of the original classes. But where we have an infinite series of classes of ascending types, this method can not be applied. For this reason, we can not now prove that there must be infinite classes. For suppose there were only $n$ individuals altogether in the universe, where $n$ is finite. There would then be $2^n$ classes of individuals, and $2^{2^n}$ classes of classes of individuals, and so on. Thus the cardinal number of terms in each type would be finite; and though these numbers would grow beyond any assigned finite number, there would be no way of adding them so as to get an infinite number. Hence we need an axiom, so it would seem, to the effect that no finite class of individuals contains all individuals; but if any one chooses to assume that the total number of individuals in the universe is (say) 10,367, there seems no à priori way of refuting his opinion.

From the above mode of reasoning, it is plain that the doctrine of types avoids all difficulties as to the greatest cardinal. There is a greatest cardinal in each type, namely the cardinal number of the whole of the type; but this is always surpassed by the cardinal number of the next type, since, if $\alpha$ is the cardinal number of one type, that of the next type is $2^\alpha$, which, as Cantor has shown, is always greater than $\alpha$. Since there is no way of adding different types, we can not speak of "the cardinal number of all objects, of whatever type," and thus there is no absolutely greatest cardinal.

If it is admitted that no finite class of individuals contains all individuals, it follows that there are classes of individuals having any finite number. Hence all finite cardinals exist as individual-cardinals; *i.e.*, as the cardinal numbers of classes of individuals. It follows that there is a class of $\aleph_0$ cardinals, namely, the class of finite cardinals. Hence $\aleph_0$ exists as the cardinal of a class of classes of classes of individuals. By forming all classes of finite cardinals, we find that $2^{\aleph_0}$ exists as the cardinal of a class of classes of classes of classes of indi-dividuals; and so we can proceed indefinitely. The existence of $\aleph_n$ for every finite value of $n$ can also be proved; but this requires the consideration of ordinals.

If, in addition to assuming that no finite class contains all individuals, we assume the multiplicative axiom (*i. e.*, the axiom that, given a set of mutually exclusive classes, none of which are null, there is at least one class consisting of one member from each class in the set), then we can prove that there is a class of individuals containing $\aleph_0$ members, so that $\aleph_0$ will exist as an individual-cardinal. This somewhat reduces the type to which we have to go in order to prove the existence-theorem for any given cardinal, but it does not give us any existence-theorem which can not be got otherwise sooner or later.

Many elementary theorems concerning cardinals require the multiplicative axiom.* It is to be observed that this axiom is equivalent to Zermelo's,† and therefore to the assumption that every class can be well-ordered.‡ These equivalent assumptions are, apparently, all incapable of proof, though the multiplicative axiom, at least, appears highly self-evident. In the absence of proof, it seems best not to assume the multiplicative axiom, but to state it as a hypothesis on every occasion on which it is used.

## X.

### *Ordinal Numbers.*

An ordinal number is a class of ordinally similar well-ordered series, *i. e.*, of relations generating such series. Ordinal similarity or *likeness* is defined as follows:

$$\mathrm{Smor} = \hat{P}\hat{Q}\,\{(\exists S)\,.\,S\varepsilon 1\to 1\,.\,\text{Ɑ}`S = C`Q\,.\,P = S\,|\,Q\,|\,\breve{S}\}\quad \text{Df.},$$

where "Smor" is short for "similar ordinally."

The class of serial relations, which we will call "Ser," is defined as follows:

$$\mathrm{Ser} = \hat{P}\,\{xPy\,.\supset_{x,y}.\sim(x=y): xPy\,.\,yPz\,.\supset_{x,y,z}.\,xPz:$$
$$x\varepsilon C`P\,.\supset_x.\,\overrightarrow{P}`x \cup \iota`x \cup \overleftarrow{P}`x = C`P\}\quad \text{Df.}$$

That is, reading $P$ as "precedes," a relation is serial if (1) no term precedes itself, (2) a predecessor of a predecessor is a predecessor, (3) if $x$ is any term in the field of the relation, then the predecessors of $x$ together with $x$ together with the successors of $x$ constitute the whole field of the relation.

---

* *Cf.* Part III of a paper by the present author, "On some Difficulties in the Theory of Transfinite Numbers and Order Types," *Proc. London Math. Soc.* Ser. II, Vol. IV, Part I.

† *Cf. loc. cit.* for a statement of Zermelo's axiom, and for the proof that this axiom implies the multiplicative axiom. The converse implication results as follows: Putting Prod $`k$ for the multiplicative class of $k$, consider

$$Z`\beta = \hat{R}\,\{(Ex)\,.\,x\varepsilon\beta\,.\,D`R = \iota`\beta\,.\,\text{Ɑ}`R = \iota`x\}\quad \text{Df.},$$

and assume

$$\gamma\varepsilon\ \mathrm{Prod}\ `Z``cl`a\,.\,R = \hat{\xi}\hat{x}\{(S)\,.\,\exists S\varepsilon\gamma\,.\,\xi Sx\}.$$

Then $R$ is a Zermelo-correlation. Hence if Prod $`Z``cl`a$ is not null, at least one Zermelo-correlation for $a$ exists.

‡ See Zermelo, "Beweis, dass jede Menge wohlgeordnet werden kann." *Math. Annalen*, Vol. LIX, pp. 514–516.

Well-ordered serial relations, which we will call $\Omega$, are defined as follows:

$$\Omega = \hat{P}\,\{P \varepsilon \text{Ser} : \alpha \subset C'P \,.\, \exists \,!\, \alpha \,.\, \supset_\alpha .\, \exists \,!\, (\alpha - \breve{P}``\alpha)\} \quad \text{Df.};$$

*i. e.*, $P$ generates a well-ordered series if $P$ is serial, and any class $\alpha$ contained in the field of $P$ and not null has a first term. (Note that $\breve{P}``\alpha$ are the terms coming after some term of $\alpha$).

If we denote by $No\,`P$ the ordinal number of a well-ordered relation $P$, and by $NO$ the class of ordinal numbers, we shall have

$$No = \hat{\alpha}\hat{P}\,\{P\varepsilon\Omega \,.\, \alpha = \overrightarrow{\text{Smor}}\,`P\} \quad \text{Df.}$$
$$NO = No``\Omega.$$

From the definition of $No$ we have

$$\vdash : P\varepsilon\,\Omega \,.\supset .\, No`P = \overrightarrow{\text{Smor}}\,`P$$
$$\vdash : \sim (P\varepsilon\,\Omega) \,.\supset .\, \sim E\,!\, No`P.$$

If we now examine our definitions with a view to their connection with the theory of types, we see, to begin with, that the definitions of "Ser" and $\Omega$ involve the *fields* of serial relations. Now the field is only significant when the relation is homogeneous; hence relations which are not homogeneous do not generate series. For example, the relation $\iota$ might be thought to generate series of ordinal number $\omega$, such as

$$x,\ \iota`x,\ \iota`\iota`x, \ldots .\ \iota^{n}`x, \ldots .,$$

and we might attempt to prove in this way the existence of $\omega$ and $\aleph_0$. But $x$ and $\iota`x$ are of different types, and therefore there is no such series according to the definition.

The ordinal number of a series of individuals is, by the above definition of $No$, a class of relations of individuals. It is therefore of a different type from any individual, and can not form part of any series in which individuals occur. Again, suppose all the finite ordinals exist as individual-ordinals; *i. e.*, as the ordinals of series of individuals. Then the finite ordinals themselves form a series whose ordinal number is $\omega$; thus $\omega$ exists as an ordinal-ordinal, *i. e.*, as the ordinal of a series of ordinals. But the type of an ordinal-ordinal is that of classes of relations of classes of relations of individuals. Thus the existence of $\omega$ has been proved in a higher type than that of the finite ordinals. Again, the cardinal number of ordinal numbers of well-ordered series that can be made out of finite ordinals is $\aleph_1$; hence $\aleph_1$ exists in the type of classes of classes of classes of relations of classes of relations of individuals. Also the ordinal numbers of well-ordered series composed of finite ordinals can be arranged in order of magnitude, and the result is a well-ordered series whose ordinal number is $\omega_1$. Hence $\omega_1$ exists as an ordinal-ordinal-ordinal. This process can be repeated any finite number of times, and thus we can establish the existence, in appropriate types, of $\aleph_n$ and $\omega_n$ for any finite value of $n$.

But the above process of generation no longer leads to any totality of *all* ordinals, because, if we take all the ordinals of any given type, there are always greater ordinals in higher types; and we can not add together a set of ordinals of which the type rises above any finite limit. Thus all the ordinals in any type can be arranged by order of magnitude in a well-ordered series, which has an ordinal number of higher type than that of the ordinals composing the series. In the new type, this new ordinal is not the greatest. In fact, there is no greatest ordinal in any type, but in every type all ordinals are less than some ordinals of higher type. It is impossible to complete the series of ordinals, since it rises to types above every assignable finite limit; thus although every segment of the series of ordinals is well-ordered, we can not say that the whole series is well-ordered, because the "whole series" is a fiction. Hence Burali-Forti's contradiction disappears.

From the last two sections it appears that, if it is allowed that the number of individuals is not finite, the existence of all Cantor's cardinal and ordinal numbers can be proved, short of $\aleph_\omega$ and $\omega_\omega$. (It is quite possible that the existence of these may also be demonstrable.) The existence of all *finite* cardinals and ordinals can be proved without assuming the existence of anything. For if the cardinal number of terms in any type is $n$, that of terms in the next type is $2^n$. Thus if there are no individuals, there will be one class (namely, the null-class), two classes of classes (namely, that containing no class and that containing the null-class), four classes of classes of classes, and generally $2^{n-1}$ classes of the $n$th order. But we can not add together terms of different types, and thus we can not in this way prove the existence of any infinite class.

We can now sum up our whole discussion. After stating some of the paradoxes of logic, we found that all of them arise from the fact that an expression referring to *all* of some collection may itself appear to denote one of the collection; as, for example, "all propositions are either true or false" appears to be itself a proposition. We decided that, where this appears to occur, we are dealing with a false totality, and that in fact nothing whatever can significantly be said about *all* of the supposed collection. In order to give effect to this decision, we explained a doctrine of *types* of variables, proceeding upon the principle that any expression which refers to *all* of some type must, if it denotes anything, denote something of a higher type than that to all of which it refers. Where *all* of some type is referred to, there is an *apparent variable* belonging to that type. Thus *any expression containing an apparent variable is of higher type than that variable.* This is the fundamental principle of the doctrine of types. A change in the manner in which the types are constructed, should it prove necessary, would leave the solution of contradictions untouched so long as this fundamental principle is observed. The method of constructing types explained above was shown to enable us to state all the fundamental definitions of mathematics, and at the same time to avoid all known contradictions. And it appeared that in practice the doctrine of types is never relevant except where existence-theorems are concerned, or where applications are to be made to some particular case.

The theory of types raises a number of difficult philosophical questions concerning its interpretation. Such questions are, however, essentially separable from the mathematical development of the theory, and, like all philosophical questions, introduce elements of uncertainty which do not belong to the theory itself. It seemed better, therefore, to state the theory without reference to philosophical questions, leaving these to be dealt with independently.