

Call Number: C0282

Location: mss

Sub Location: ""

Volume/Box Number(s): Box 7b ,
Folder(s): 30

Author: Gödel, Kurt, 1906-1978.

Collection: Kurt Gödel Papers

Collection is open for research use.

Finding aid ID's: C0282_c00686

"The existence of undecidable propositions in any formal system containing arithmetic,"
(lecture to Philosophical Society of N.Y. Univ., 18 Apr. 1934) : manuscript text and typed abstract

04/30

I appreciate very much my having the opportunity of speaking before you on some of the modern results in the foundations of math. and I hope I shall succeed in making ~~the~~ leading ideas clear to you, despite the very technical character of the work. The subject I want to talk about is closely connected with the so-called formalisation of math. i.e. with the fact ~~discovered~~ that all math. and logic (at least all math. and logic that has been developed so far) can be deduced by means of a few axioms and rules.

The modern invest. in the found. of math. gave as one of their outstanding results the fact, that all math. and logic (at least all math. that has been developed so far) can be deduced by means of a few axioms and rules of inf. **T** In order to bring out this fact clearly it was necessary at first to replace the imprecise and often ambiguous colloquial language (in which math. statements are usually expressed) by a perfectly precise artificial language. This is called logistic formalism. This formalism consists of a few prim. symb. which represent the prim. notions of log. and math. and ^{play the same role as} ~~correspond to~~ the words in ordinary language. I wrote some examples of the prim. symbols on the blackboard. —

Now any log. or math. prop. can be expressed by a formula composed of these prim. symbols and vice versa any formula composed of our

prim terms according to certain rules [which constitute the grammar of our log. language] expresses a definite math. statement. In practice it would be very inconvenient to express math. statement in this way by means of the prim terms of. ~~itself~~ i.e. our formulas would become very long and cumbersome. Therefore ~~we usually introd.~~ ^{are introduced} besides our prim terms new symbols by def. but it is to be noted that this [device] serves merely the practical purpose of abbrev. and is therefore entirely dispensable from the theor. pt. of view. Since we can replace in every formula the new symbols by their meaning expr. in ~~the~~ ^{the} prim terms. So we may disregard the possibility of introd. new symbols by def. and think of any math. statement as expressed by our prim terms alone. — The process of deduction i.e. of proof is represented in our formalism in the

following manner: Some of our formulas are considered as axioms, i.e. as the starting point for development and in addition to that certain rules of inf are stated which allow one to pass from the axioms to new formula's and thus ded. more - prop. One of the rules of inf. e.g. reads. If A and B are two arbitrary formulas and if you have proved the formula A and $A \rightarrow B$ you are entitled to conclude B . The other rules of inf. are of a similar simple character. In partic. all of them are purely formal i.e. they do not refer to the meaning of ~~the~~ ^{the} formulas but only to their outward structure ^{and} so that they could be applied by some one who knows nothing about the meaning of the symb. One could even ^{possibly} device a machine which would give you as many ^{correct} ^{of the axioms} consequences as you like; the only trouble

would be that it - at random and therefore
not the results one is interested in. By iterated
application of the rules of inf. starting from the
axioms we obtain what I call a chain of inference.

By A. ch. of inf. is simply a finite sequence of
formula's $A_1, \dots, A_n$ which begins with some of our axioms
and has the prop. that each of it's other formula's
can ^{be} obtained from some of the preceding one's
by ^{an} appl. of ^{one of} our rules of inf. Instead of chain of
inf. I shall also use the term formal proof or
briefly proof. ~~That~~ A proof ending up with the formula
 F is called a proof for the formula F

~~(the last formula in this proof.)~~ and of course we
shall call a formula F provable if there is a proof
for it; which means the same thing as: F can
be obtained from the axioms by iterated application
of the rules of inference. A symbolism for which

axioms and rules of inf. are specified in the manner I have just described is called a formal system and the fact to which I referred in the beg. of my talk ^{can now be expressed by saying} ~~that~~ that one has succeeded in reducing all of ^{math and log. to} ~~math and log.~~ ^{in such a way that every math proof can be} ~~to~~ a formal system. ^{Due} to this fact certain general questions concerning the structure of math. which formerly had to be left to vague speculations [and could not even be stated precisely], have become amenable to scientific treatment. I want to deal with ^{of these} two ~~such~~ questions. The first concern's the freedom from contradiction of math. This question can now be stated in a perfectly precise way as follows.

"Does there exist any formula A such that A and not $\sim(A)$ are both provable" where the term provable has the precise meaning which I defined before namely it means "obtainable from the axioms by our rules of inf. in a finite number of steps."

It can easily be shown that if there existed two formula's A and $\neg A$, both of which were provable, then any formula whatsoever would be provable for instance also the formula $0=1$. So it is of vital importance for our formalism that this should not happen and the problem of giving a proof that it cannot happen arises. But at the same time an ~~serious~~ ^{can be brought} objection against the soundness of this problem. Namely one may say: Suppose we had given a proof for consistency then owing to the fact that it is a math. proof it must necessarily proceed according to the axioms and rules of inf. for math. and log. So in order to be convinced by this supposed proof we must know that our axioms and rules of inf. which we used allways lead to correct results. But if we know this in advance then no proof for freedom from contradiction is necessary

[because rules of inf. which lead to correct results cannot lead to A and $\sim A$ because these two formulas cannot both be correct] Fortunately the actual situation is slightly different. For math. consists of two distinct parts which are usually referred to as finite and transfinite math. and which may be roughly characterised as follows. Under the first heading (of f. math.) are comprised all such methods of proof which do not presuppose the existence of any infinite set whereas under the second heading (of t. math.) fall those math. of proof which do presuppose the existence of infinite sets and are based on this assumption. [e.g. let P be any arithm. prop. and let's consider the statement either every int. has the prop. P or there is an integer which has not this prop.]

Now nobody has ever questioned seriously the consistency of finite math. whereas the situation is quite different with the transfinite methods based on the assumption of the existence of inf. sets, which by the way is by far the greater part of math. now existing. In this domain of math. actual contradictions had arisen unexpectedly toward the end of the 19. century the so called paradoxes of the theory of aggreg. In order to avoid them certain restrictions on the previous assumptions concerning ^{existence of} the infinite sets had to be made. These restrictions can be made in a very natural way and they do not ^{affect in any way} ~~alter anything~~ in the math. ^{previously obtained} results, but nevertheless the faith of many math. in the transfinite methods

was shaken by this bad experience and there
 remains the fear that other paradoxes may arise
 in spite of ~~one~~ ^{the} restrictions. Now I think
 it is clear what the question of proving freedom
 from contradiction really is about. It is the
 problem of proving the freedom from contradiction
 of ~~the~~ transfinite math. by means of finite methods
 i.e. using in the proof for consistency only such
 methods as are not based on the existence of in-
 finite sets. So much for the meaning of the 1. problem,
 the question of consist. The second problem is ^{in its treatment} so closely
 related to the first that it can hardly be dealt with
 separately. It the question of completeness of ~~the~~ ^{the}
 formal system for math. i.e. the question whether
 every math. statement expr. by a formula of the
 system can be decided (either in the aff. or in the neg.)
 by means of ~~the~~ ^{the} rules of inf. and axioms, i.e. Is it

expressing a prop.

True that if A is any arbitrary formula then either A or not $\sim A$ is provable? or are there formula's for which neither ^{one} of the two is provable? I am going to sketch a proof which answers both questions in the neg. in the following sense 1) It is not possible to prove our system consistent, using only a part of the meth. of proof embodied in its axioms and rules of inf. in fact it is not even possible to prove it consistent using all of it's methods of proof 2) There are prop. in fact even arithmetic prop. ^(belonging to the system) which cannot be decided by a formal proof.

Of course math. can be formalised in different ways i.e. the ax. and rules of inf. representing math. can be chosen in different manners and so one may suspect that our two results depend on the special system for math. which we ~~are~~ choose. But this is not the case. It can be shown that

the two theorems which I just stated hold good whatever formal system we may choose provided only that ~~ordinary~~ arithmetic of integers in its usual form is contained in the system and that no false arithmetic statement is provable ...i.e. the axioms and rules of inf should not lead to results which can be disproved for intuitive reasons.

The proof for these two statements (imposs. of a proof for consistency and existence of model prop.) is very cumbersome if worked out in all details but ~~nevertheless I think one can make~~ ^{possibly make} its leading ideas clear in one lecture and I ~~shall~~ ^{shall} proceed in doing so. but I hope to succeed in making clear to the leading ideas clear to you

Suppose system given

Among expressions also such as $x > 6$

not prop. but become so if subset

called prop. function

Express properties

Similarly with several variables expressing rel.

Let ^{in interpretation the} prim. symbols be $\sim, \rightarrow, E, X, \wedge, \vee, \dots$ Now

Any formula = ^{combination} ~~sequence~~ of prim. symbols = sequence

Therefore numbering possible

In many ways, we choose the following:

[# number prim. symbols --]

Proof = sequence of formulas = sequence of numbers

Numbering of proofs

Not all numbers used up but one to one

Owing to numbering: class of formula's $\Rightarrow$ class of numbers
relation " $\Rightarrow$ relation " "

e.g. relation of being longer

Similarly for any relation (called metamath) $\Rightarrow$

relation arithmetic

1202
040123

A relation between formulae such as being longer is
Analogous to analytical geom. (also statements)

Further examples needed for sub. proof.

Relation of im. consequence $\frac{P}{Q}$
 R

means $P \xrightarrow{\text{the formulae}} Q \rightarrow R$ (=impl with Q as
first and R as second term)

What does that mean for con. numbers p, q, r

Series of exp con. to series of symbols

therefor series of exp of p must be comp of those for
 q and r with one between them

purely arithm. relation

call it for the moment derived

for any three numbers it can be ascertained whether or not

Arithmetic definition of the integers which are
numbers of proof as follows:

Suppose n Axioms with numbers $k_1, \dots, k_n$
(definite numbers which can be computed)

Suppose further only one rule of inf which
makes no essential difference

Recall def. of formal proof (or chain of inference)

What does that mean for numbers?

According to correspondence if n is number of proof then exponents numbers of formulas occurring in this proof and so the exp must satisfy this condition 1) first - $k_1 - k_n$ 2. each derived from some of preceding one where derived means

This property again purely arithmetic proof numbers

Further we consider the relation $\boxed{x \text{ Pr } y}$

means x Proof and last exp. of $x - y$

and class $P(x) \equiv (\exists y) y \text{ Pr } x$

^{since} ~~the~~ the notions $\boxed{P}$ and $\boxed{\text{Pr}}$ are arithmetic and as arithm. contain in our system they can be expr. by formulas in fact by prop. functions as can be shown in detail.

So consider from now on $\boxed{P}$ and $\boxed{\text{Pr}}$ as abbreviations for complicated formulas which can actually be found and written

Rel. $\exists Pn x$ constructive (finite number of steps) and this has the consequence.

If A arbitrary formula and a its number
 If A provable then $P(a)$ provable

Proof sup. A provable and b number
 then $b Pn a$ true and provable

hence $(\exists y) \exists Pn a = P(a)$ provable

I need one more arithm. notion derived from metarith. $S(x, y)$ (calculable)

Again can be shown to be arithmetic notion
which can be calculated (how?)

Therefore represented by a formula of our System and again consider S as denoting this formula

through with preparations

Consider this expression

$$\sim P[S(x, x)] \quad \dots \quad [9]$$

This is a prop. function with one variable and means:

1. formula obtained by subst x in formula number x is not provable
2. ~~(x is not prop. function then always true)~~
2. The property expressed by prop. function number x cannot be proved to belong to the number x - computed

The above prop. f. being a formula of our system it must have a number q (calculable!)

Subst q y get a prop.

$$\sim P[S(qq)] \quad S(qq)$$

which says that prop. number $S(qq)$ is not provable. What is the number of this formula $S(qq)$.

Lets introduce n for $S(qq)$ then

$\underbrace{\sim P(a)}_A \quad \text{---} \quad a$
number

A states on itself that it is not provable
or is arithmetic statement equivalent to
statement A not provable

Now we prove

51 A provable then system contradictory

Apply auxiliary theorem we have a prop.

A with number d and know if A provable

then $P(a)$ provable in

$\neg \sim P(a)$ provable then $P(a)$ provable

$\mathcal{I} \sim P(u)$ provable system contradictory

If system consistent A not provable

But owing to the fact that A itself means exactly that A is not provable we may say

If system consistent then A i.e.

$$C \rightarrow A$$

if C means the statement that the system

is not provable.

This statement can itself be expressed by formula owing to correspondence

So we proved a certain formula of our system $C \rightarrow A$ and this proof can be formalised so we have

$(C \rightarrow A)$ is provable

Now it follows that C cannot be proved because if it were provable then A were provable and then system contradictory

~~these considerations show consistency~~

~~we could also prove $\neg P(a)$ and we know if
we prove $\neg P(a)$ then we could also prove
 $P(a)$ and the system is contradictory.~~

So we have shown if the statement that our system is free from contradiction could be proved then our system would be contradictory and a closer examination shows that we could actually exhibit this contradiction i.e. given a proof for freedom from contradiction we could ~~transform~~ derive from it an actual contradiction of our system. The second half of our program ^(proof for the) existence

of undecidable ^{arithm} propositions is now easily accomplished
e.g. A is such an undecidable prop. For ^{we know if} ~~suppose~~ that our system is ~~correct~~ free from contradiction then A is not provable
on the other hand ^{under the same assumption} we know that A is true therefore
 $\neg A$ is false and therefore $\neg A$ cannot be provable
~~then~~ ^{assume} that no false arithmetic statement
is provable in our system. ~~At the same time~~

The prop. A which we proved to be undecidable
~~seems to be very artificial and far remote from~~
 is an arithmetic statement because P ~~is an arithmetic~~
~~property~~ and S ~~consists~~ of which it is ~~built~~
 constructed are arithmetic notions. But ^(this prop. at first sight) it seems to
 be very artificial and far remote from everything
 that is actually dealt with in arithmetic. This is
 however is a ^{strong} ~~false~~ appearance, ^{I can be shown that} ~~since it~~
 can be transformed into a statement on the
 solutions of a certain diophantine ^{equation}, ~~therefore~~
~~this statement too, must be undecidable~~
 i.e. into a statement of the same character as are
 actually dealt with in number theory.

Of course ~~the~~ ^{the} undecidability of A is only relative. We can add a new axiom to our system which has the consequence that A becomes decidable in fact a very plausible axiom namely C which states that our system is free from contradiction. If we add this C , then owing to this implication A becomes provable but it would be wrong to suppose that now we should have obtained a system in which every arithmetic statement is decidable. For we can apply the same method of proof ~~to~~ to our new system and construct ^{another} ~~prop.~~ prop., which is undecidable in the new system ^{and} and so we can go on indefinitely without ever reaching a system in which every arithmetic statement is decidable. This situation can also be expressed by saying it is impossible to give a complete system of axioms for ^{the} arithmetic of integers i.e. a system

^{system is}
 $\forall$ For this system again we can add a new axiom e.g. the statement that the new and make the ~~and~~ by it the undecidable prop. a decidable one

which ~~would~~ makes it possible to decide any given ^{with} statement expressible in the prim. terms of our system. Finally I wish to make a final remark on the impossibility of proving consistency. I think even for our statement is only relative i.e. we proved only that if a definite formalization of math. is given then it is impossible to prove consistency of that formal system ~~by~~ i.e. using ~~the~~ only the axioms and rules of inf. of this same system. Some one may set up another formalism of math. and prove the consistency of the first system by an argument proceeding according to the rules of the second system. But we know in a ~~shortly~~ proof for consistency the point is that it should be conducted by finite methods and now nobody has ever been able to produce a proof conducted by finite methods which could not easily be expressed in any ^{one} ~~one~~ of the ~~known~~

formal systems for mathematics and nobody knows
 how to construct such a proof and therefore the
 foregoing ^{considerations} ~~considerations~~ make it
 appear entirely hopeless to prove consistency for
~~finite arithmetic~~ the transfinite methods of
 math. using only the finite unobjectionable methods
 of finite arithmetic which was the program
 of the formalist school.

NOTE

The following pages comprise a translation of the less technical portion of Dr. Kurt Gödel's paper "Über formal unentscheidbar Sätze der Principia Mathematica und verwandter System I", Monatshefte, vol. 38 (1931), pp. 173-198. Dr. Gödel is to lecture to the New York University Philosophical Society on April 18; and this translation is supplied so that, in case you have not the opportunity of consulting the original article, you may nevertheless acquaint yourself with his more general ideas. Another paper by Dr. Gödel on mathematical logic is "Die Vollständigkeit der Axiome des logischen Funktionenkalküls", Monatshefte, Vol. 37 (1930), pp. 348-360.

Since there exist no precise English equivalents for the German words "beweisbar", "entscheidbar", "unbeweisbar", and "unentscheidbar", we have seen fit to let these words stand in German wherever they occur in the text. A proposition is said to be "beweisbar" in a system if it can be proved within that system; if it cannot be proved within the system it is "unbeweisbar". If the contradictory of a proposition can be proved within a system then the contradictory of the proposition is beweisbar within the system; and then the proposition itself is (assuming the system consistent) unbeweisbar. Thus in the system of Euclidean geometry the proposition, "The sum of the interior angles of a triangle equals two right angles" is beweisbar, and the proposition, "The sum of the interior angles of a triangle is less than two right angles" is unbeweisbar. (It is obvious that all beweisbar propositions are true within the system, but a true proposition is not necessarily beweisbar.) If it happens that both a proposition and its contradictory are unbeweisbar within a system, then that

proposition is said to be "unentscheidbar" within the system. Thus in Euclidean geometry the proposition "Plato preceded Aristotle" is unentscheidbar; for the ideas and axioms of Euclidean geometry do not suffice to permit us to prove either that Plato preceded Aristotle or that it is false that Plato preceded Aristotle. Propositions which are not unentscheidbar are said to be entscheidbar.

UPON FORMAL UNENTSCHEIDBAR PROPOSITIONS OF "PRINCIPIA
MATHEMATICA" AND OF RELATED SYSTEMS

By Kurt Gödel

The development of mathematics in the direction of greater rigor has, as is familiar, led to the point where wide branches of mathematics have been formalized in such a way that proofs can be carried out according to a few mechanical rules. At the present time the most comprehensive established formal systems are Principia Mathematica and the Zermelo-Fraenkel axiom system. These two systems are so general that all methods of proof used today in mathematics are formalized in them; that is, the methods of proof are reduced to some few axioms and rules for deduction. Therefore the assumption might seem reasonable that these axioms and rules of deduction suffice for this much, that all mathematical questions which permit formal expression in the systems in question are also entscheidbar. In the following it is shown that this is not the case but that there are in the two above-mentioned systems even relatively simple problems in the theory of the ordinary whole numbers which are unentscheidbar. This

circumstance arises, not from the special nature of the systems established, but obtains in a very wide class of formal systems to which belong especially those which arise from the two systems mentioned by the addition of a finite number of axioms. V

We sketch here the main ideas of the proofs, naturally without claiming to reach any great exactness. The laws of a formal system (we limit ourselves here to the system of Principia Mathematica) are, formally considered, finite rows of fundamental marks (variable, logical constant, parentheses or punctuation dots) and one can easily decide which rows of fundamental marks are meaningful combinations of symbols and which are not. Analogously proofs, considered from the formal standpoint, are nothing more than finite rows of marks (with determined storable properties). One can show that the concepts, "propositions", "forms of proof", "beweisbar propositions", are definable within the Principia Mathematica system; that is, one can exhibit, for example, a proposition $F(v)$ from the Principia Mathematica with a real variable v of the type of a number sequence, so that $F(v)$, when meaningfully interpreted, says " v is a beweisbar proposition". < 0

We now present an unentscheidbar proposition of the Principia Mathematica; that is, a proposition A , such that neither A nor $\text{non-}A$ is beweisbar.

An expression from the Principia Mathematica with exactly one ^{real} variable and the latter of the type of the natural numbers (class of classes) we shall call a class symbol. The class symbols we regard as ordered in some way in a sequence* and we denote the

*Perhaps according to increasing numbers of letters in the symbols; and, when the numbers are the same, alphabetically.

nth one by $R(n)$. We remark that the concept "class symbol" as well as the ordering relation R can be defined in the Principia Mathematica system. Let α be any class symbol; by $[\alpha; n]$ we denote that proposition which arises from the class symbol α when one replaces the real variable by the symbol for the natural number n . Also the triadic relation $x = [y; z]$ is evidently definable within the Principia Mathematica system. We now define a certain class K of natural numbers in the following way:

$$n \in K = \overline{\text{Bew}} [R(n); n] \quad \text{Def.} \quad (1)$$

(where $\text{Bew}(x)$ indicates: "x is a beweisbar proposition", and the upper bar indicates negation). Since the concepts which appear in the definiens are all definable in the Principia Mathematica, so also is the concept K , set up from them--that is, there is a class symbol S such that the expression $[S; n]$, when non-symbolically interpreted, asserts that the natural number n belongs to the class K . S is identical as a class-symbol with a definite $R(q)$; that is, it is the case that

$$S = R(q)$$

for a definite natural number q . We show now that the proposition $[R(q); q]$ is unentscheidbar in the Principia Mathematica system. For suppose the proposition $[R(q); q]$ were beweisbar; then it would be true also; then, according to the above, q would belong to K --that is, according to (1) it would be true that $\overline{\text{Bew}}[R(q); q]$ holds contrary to hypothesis. On the other hand, were the negation of $[R(q); q]$ beweisbar, then it would be true that $n \in K$, or $\text{Bew}[R(q); q]$ holds. $[R(q); q]$ would therefore, together with its negation, be beweisbar, which again is impossible.

The proof-method just presented may manifestly be applied to any formal system which, first, non-symbolically interpreted, provides sufficient means of expression to define the concepts appearing in the above discussion (especially the concept "beweisbar formula") and in which, secondly, every beweisbar formula is intuitionally seen to be correct.

From the remark that $[R(q);q]$ affirms its own unbeweisbarkeit it follows immediately that $[R(q);q]$ is correct, for $[R(q);q]$ is unbeweisbar (because unentscheidbar). The unentscheidbar theorem in the system of Principia Mathematica is decided then by meta-mathematical considerations. Exact analysis then of this remarkable circumstance leads to some surprising results with respect to the consistency proofs of formal systems which are more closely examined in a latter section.