Kurt Gödel

"The existence of undecidable propositions in any formal system containing arithmetic"

Gödel's notes for his lecture to the Philosophical Society of New York University, 18 April 1934

Box 7b, folder 30, the Kurt Gödel Papers, the Shelby White and Leon Levy Archives Center, Institute for Advanced Study, Princeton, NJ, USA, on deposit at Princeton University Library

Text transcribed by Stephen Budiansky

0

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 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.

1

The modern investigations in the foundations of math. gave as one of the outstanding results the fact, that all math. and logic (at least all the math. that has been developed so far) can be deduced by means of a few axioms and rules of inference.

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 the logistic formalism. This formalism consists of a few prim[itive] symb[ols] which represent the prim[itive] notions of log. and math. and play the same role as the words in ordinary language. I wrote some examples of the primitive symbols on the blackboard. Now any log. or math. prop[erty] can be expressed by a formula composed of these prim. symbols and vice versa any formula composed of our

2

prim. terms according to certain rules (which constitute the grammar of our log. language) expresses a definite mathematical statement. In practice it would be very incon[venient] to expr[ess] math. statements in this way by means of the prim. terms of . . . i.e. our formulas would become very long and cumbersome. Therefore besides our prim. terms new symbols are introduced by def[inition] but it is to be noted that this device serves merely the practical purpose of abbrev. and therefore it is entirely dispensable from the theor. p. o. view since one can replace in every formula the new symbols by their meaning expr. in the prim. terms. So we may disregard the possibility of introducing new symbols by def. and think of any mathematical statement as expressed by our prim. terms alone.

The process of deduction, i.e. of proof is represented in our formalism in the

3

following manner: Some of our formulas are considered as axioms, i.e. as the starting point for devel[oping] math, and in addition to that certain rules of inf[erence] are stated which allow one to pass from the axioms to new formulas and thus ded[uce] more prop[ositions]. 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 $A \to B$ you are entitled to conclude B. The other rules of inf. are of a similar simple character. In practice all of them are purely formal, i.e. they do not refer to the meaning of the formulas but only to their outward structure and so they could be applied by someone who knew nothing about the meaning of the symbols. One could even easily devise a machine which would give you as many correct consequences of the axioms as you like, the only trouble

4

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. A ch. of inf. is simply a finite sequence of formulas $A_1 \ldots A_n$ which begins with some of our axioms and has the prop[erty] that each of its other formulas can be obtained from some of the preceding ones by an application of one of our rules of inf. Instead of chain of inf. I shall also use the term formal proof or briefly proof. A proof ending up with the formula F is called a proof for the formula F 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

5

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 one has succeeded in reducing all of math. and log. to a formal system in such a way that every mathematical proof can be [ ]. Owing 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.

The first concerns 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 ~ (A) are both provable” where the term provable has the meaning which I defined before namely it means “obtainable from the axioms by our rules of inference in a finite number of steps.”

6

It can easily be shown that if there existed two formulas A and ~ 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 problems of giving a proof that it cannot happen arises.

But at the same time an objection can be brought 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 mathematical proof it must necessarily proceed according to the axioms and rules of inference for math. and logic. So in order to be convinced by this supposed proof we must know that our axioms and rules of inference which we used will always lead to correct results. But if we know this in advance then no proof for freedom from contradiction is necessary

7

(because rule of inference which lead to correct results cannot lead to A and ~ A because these two formulas cannot both be true).

Fortunately the actual situation is slightly different. For mathematics 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 finite math.) are comprised all such methods of proof which do not presuppose the existence of any infinite set whereas under the second heading (of transfinite mathematics) fall those methods of proof which do presuppose the existence of infinite sets and are based on this assumption. (e.g. let P be any arithmetic prop[erty] and let’s consider the statement either every integer has the prop. P or there is an integer which has this prop.)

8

Now nobody has ever questioned seriously the consistency of finite mathematics 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 mathematics now existing. In this domain of mathematics actual contradictions had arisen unexpectedly by toward the end of the 19. century the so called paradoxes of the theory of aggreg[atio]n. In order to avoid them certain restrictions on the previous assumptions concerning the existence of infinite sets had to be made. These restrictions can be made in a very natural way and they do not affect in any way the mathematical results previously obtained, but nevertheless the faith of many math. in the transfinite methods

9

was shaken by this bad experience and there remains the fear that other paradoxes may arise in spite of 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 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 infinite sets. So much for the meaning of the 1. problem, the question of consistency. The second problem is in its treatment so closely related to the first that it can hardly be dealt with separately. It is the question of completeness of the formal system for math. i.e. the question whether every math. statement expressed by a formula of the system can de decided (either in the affirmative or in the negative) by means of the rules of inference and axioms i.e. Is it

10

true that if A is any arbitrary formula expressing a prop. then either A or not ~ A is provable? or are there formulas for which neither one of the two is provable? I am going to sketch a proof which answers both questions in the negative in the following sense.

1.) It is not possible to prove our system consistent, using only a part of the method of proof embodied in its axioms and rules of inf. in fact it is not possible to prove it consistent using all of its methods of proof

2.) There are prop[ositions] in fact even prop. belonging to the [ ] 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. we choose. But this is not the case. It can be shown that

11

the two theorems which I just stated hold good whatever formal system we may choose provided only that 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 intuitif reasons.

The proof for these two statements (imposs. of a proof for consistency and existence of undec. prop.) is very cumbersome if worked out in all details but I hope to succeed in making the leading ideas clear to you.

12

Suppose system given
Among expressions also such as $x_2 > 6$
not prop. but becomes so if subst
called prop. function

express properties
similarly with several variables expressing rel.

————————————————————————

Let whatever the prim. symbols be ~, →, E, s5, . . . sn
Any formula = combination of prim. symbols = sequence
Therefor 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 but one to one

—————————————————————————

Owing to numbering: class of formulas ⇒ class of numbers
relation   ″   ⇒ relation ″ ″

e.g. relation of being longer
Similarly for any relation (called metamath) ⇒
relation arithmetic

13

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

—————————————————————————

Further examples needed for subs. proof.

Relation of imm. consequence $\begin{matrix} P \\ \underline{Q.} \\ R \end{matrix}$

means P the formula Q → R (= impl Q as
first and R as second term)
What does that mean for corr. numbers p, q, r
Series of exp cor. to series of symbols
therefor series of exp. of p must be comp to those for
q and r with one between them
purely arithm. relation
call if 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 $k_1 \ldots k_n$
(definite number which can be computed)
Suppose further only one rule of inf which
makes no essential difference

14

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 preceding one where derived means

— — —

This property again purely arithmetic proof number
Further we consider the relation y Pr x
means x Proof and last exp. of x = y

— — —

and class P(x) ≡ (E y) y Pr x

———————————————————————————————

since the notions P and Pr are arithmetic and as arith. contain in our system they can be exp. by formulas in fact by prop. functions as can be shown in detail.

———————————————————————————————

So consider from now on P and Pr
as abbreviations for complicated formulas
which can actually be found and written

15

Rel. y Pr x constructive (finite number of steps) and this has the consequ.

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

Proof sup A provable and b number
then b Pr a true and provable
hence (Ey) y Pr a = P(a) provable

—————————————————————————————————

I need one more arith. notion derived
from metamat. 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

16

Consider this expression

$$\sim P\,[\,S(x, x)\,] \;.\;.\;.\;\; q$$

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

1. formula obtained by subst. x in formula number x is not provable
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 I get a prop.

$$\sim P\,[\,S(q\;q)\,] \qquad\qquad S(q\;q)$$

which says that prop. number S(q q) is not provable. What is the number of this formula S(q q).
Lets introduce a for S(q q) then

17

$$\underbrace{\sim P(a)}_{A} \qquad\qquad \underbrace{a}_{\text{number}}$$

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

— — — — — — —

Now we prove
If A provable then system contradictory
Apply auxiliary theorem we have a prop. A with number a and know if A provable then P(a) provable so

If ~ P(a) provable then P(a) provable
If ~ P(a) provable system contradictory
If system consistent <u>A not provable</u>
But owing to the fact that A itself means exactly that A is not provable we may say
<u>If system consistent then A</u> i.e.

$$C \rightarrow A$$

if C means the statement that the system

18

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 the system contradictory

19

So we have shown if the statement that out 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 of contradiction we could derive from it an actual contradiction of our system. The second half of our program the proof of the existence of undecidable propositions is now easily accomplished e.g. A is such an undecidable prop. For we know if our system is free from contradiction then A is not provable.

19.1

The prop. A which we proved to be undecidable is an arithmetic statement because P and S of which it is constructed are arithmetic notions. But this prop. A seems at first sight to be very artificial and far remote from everything that is actually dealt with in arithmetic. This however is a wrong appearance. It can be shown that A can be transformed into a statement on the solutions of a certain Diophantine equation, i.e. into a statement of the same character as one actually dealt with in number theory

20

Of course 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 our new system and construct another prop., which is undecidable in the new system, 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

21

which makes it possible to decide any given arith. statement expressible in the prim. terms of our system.

I wish to make a final remark on the impossibility of proving consistency. In that case too our statement is only relative i.e. we proved only that if a definite formalisation of math. is given then it is impossible to prove consistency of that formal system i.e. using only the axioms and rules of inf. of this same system. Someone 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 proof of 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 of the

22

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