THE GÖDEL THEOREM

An Informal Exposition

NORWOOD RUSSELL HANSON

Gödel demonstrates that any logical system which includes Arithmetic must be incomplete. For within such a system there will always be (well-formed, meaningful) formulae $\mathfrak{F}$, which are 'undecidable',—such that neither $\mathfrak{F}$ nor $\sim \mathfrak{F}$ is a theorem. Thus no decision procedure exists for Arithmetic; indeed if one *did* exist Arithmetic would be contradictory, — this is the crux of Gödel's proof. Hence, if Arithmetic is consistent it must be incomplete.

The original proof of this[1] is very difficult. Most informal expositions, however, convey too little of the power and ingenuity of Gödel's argument. Perhaps we can steer a middle course. Our aim will be intelligibility without undue sacrifice of rigour. The actual deduction (IV) presupposes a prior discussion of I. Decision Procedure, II. Recursive Functions and III. The Arithmetization of Logical Syntax.

I. DECISION PROCEDURE

General Notions:

A *formal mathematical system* consists of symbols, and rules for their manipulation.

Primitive (undefined) terms are the individual symbols.

Formulae are finite sequences of primitive terms.

Meaningful formulae are symbol-sequences constructed according to the rules of the system.

Axioms (primitive formulae) are a sub-class of all the meaningful formulae.

A rule of inference (R) defines the relation of 'immediate consequences by R' between a set of meaningful formulae (premises) and a further meaningful formula (conclusion). This R may be 'from $\mathfrak{A}$ and $\mathfrak{A} \supset \mathfrak{B}$ infer $\mathfrak{B}$ '.

A finite procedure must be available for determining whether a formula is meaningful, or whether a conclusion is an immediate consequence of a set of premises.

Received May 16, 1960

A proof consists in a finite sequence of meaningful formulae. Each of these formulae is either an axiom or an immediate consequence of preceding formulae. The last member of this sequence is the theorem proved.

A provable formula (a theorem) is a meaningful formula for which a proof exists.

The formal system is *complete* if for any meaningful formula $\mathfrak{F}$, either $\mathfrak{F}$ or $\sim \mathfrak{F}$ is a theorem.

'Effective': Ideally, a formal system should provide an *effective* criterion for the recognition of genuine formulae, proofs and theorems. From direct inspection (and from following fixed directions as to symbol-manipulation), it should be possible to determine 1) whether a finite sequence of primitive terms is a (meaningful) formula, 2) whether a finite sequence of formulae is a proof, and 3) whether a certain formula is a theorem, i.e. a strict consequence of a given set of formulae. The test of whether a finite sequence of formulae is a proof of the final formula is *effective*. We can always determine whether a formula is primitive (an axiom), or is a theorem. Direct inspection (and the following of rules as to symbol-manipulation) is sufficient for this. If it were not, no 'proof' would be compelling; one could always challenge a proof by asking for a proof that it was a proof, *ad indefinitum*. However, the test of whether a given formula is a theorem (by the criterion above, that it is a theorem if a proof of it exists), this test is *not in general effective*. Failure to find a proof of $\mathfrak{F}$ may derive from a lack of ingenuity rather than the actual non-existence of a proof.

Decision Procedure: The problem of supplying an effective test by means of which it can always be determined whether a formula (of a certain system) *is a theorem*, constitutes the *decision problem* for that system. Suppose that 'the system' is the propositional calculus of *Principia Mathematica.*[2] The decision problem here is usually solved by the *truth-table decision procedure*. Thus, given a formula of the calcul s, and the assignment of a truth-value to each of its variables, one can determine by a purely mechanical process the truth-value of the entire formula. This procedure does not depend on ingenuity (as does the actual deduction of a given formula from a set of primitive formulae). One has only to apply the truth-table rules; the decision whether a given formula is or is not a theorem is made automatically.[3] If for all possible assignments of truth-values, the calculated value for the entire formula is **T** (truth), the formula is said to be a tautology. The test of whether a formula is a tautology is an *effective test*; the number of possible alternative assignments of truth-values to variables is finite. The truth-value of the entire formula can be calculated for every assignment of truth-values to the variables. Hence, the test whether a formula is a tautology provides a solution to the decision problem of the propositional calculus. [It will appear that there is no such solution to the decision problem of Arithmetic.]

II. RECURSIVE FUNCTIONS

General Notions:

$x, y, z, \ldots$	non-negative integers
$\pi, \phi, \psi, \ldots$	functions of one or more integers
n	some particular x, y or z
$\mathbf{S}$	the successor function;
$(\perp y)\ (\exists x)$	'. . . for all y there is an x such that . . .'
$\mu\ k$	'the smallest k such that . . .'

'*Recursive*': The simplest example of proof by recursion is the proof that every non-negative integer has some property π; 1) that 0 has π and 2) that if x has π then $(x + 1)$ also has π.

'*Primitive Recursive*':

1. A (monadic) function π is defined by primitive recursion in terms of a (dyadic) function ϕ by the equations:

$$\pi\ (0) = n$$

$$\pi\ (\mathbf{S}\ (x)) = \phi\ (x, \pi\ (x))$$

2. And a (dyadic) ϕ may be defined by primitive recursion in terms of a (monadic) π, and a (triadic) ψ, by the equations:

$$\phi\ (x, 0) = \pi\ (x)$$

$$\phi\ (x, \mathbf{S}\ (y)) = \psi\ (x, y, \phi\ (x, y))$$

which hold for all, x and y.

3. Consider e.g., Peano's definition of addition:

$$x + 0 = x$$

$$x + \mathbf{S}\ (y) = \mathbf{S}\ (x + y)$$

This comes under the general form of definition by primitive recursion, where $\pi\ (x) = x$, and where $\psi\ (x, y, z) = \mathbf{S}\ (z)$.

4. Another example is Peano's introduction of multiplication by the pair of equations:

$$x \times 0 = 0$$

$$x \times \mathbf{S}\ (y) = (x \times y) + x$$

Addition is taken as previously defined, and $\pi\ (x) = 0$, while $\psi\ (x, y, z) = z + x$.

5. In general a function from non-negative integers to non-negative integers is *primitive recursive* if it can be obtained according to the following scheme:

i) The function $\phi\ (x_1 \ldots x_n)$ is *composite* with respect to $\psi\ (x_1 \ldots x_n)$ and $\pi_i\ (x_1 \ldots x_n)$, – where $i = 1, \ldots n$, – if, for all natural numbers $x_1 \ldots x_n$,

$$\phi\,(x_1 \ldots x_n) = \psi\,(\pi_1\,(x_1 \ldots x_n) \ldots \pi_n\cdot(x_1 \ldots x_n))$$

ii) The function $\phi\,(x_1 \ldots x_n)$ is *recursive* with respect to $\psi\,(x_1 \ldots x_{n-1})$ and $\pi\,(x_1 \ldots x_{n+1})$ if, for all natural numbers $k, x_2 \ldots x_n$,

$$\phi\,(0, x_2 \ldots x_n) = \psi\,(x_2 \ldots x_n)$$

$$\phi\,(k+1, x_2 \ldots x_n) = \pi\,(k, \phi\,(k, x_2 \ldots x_n), x_2 \ldots x_n)$$

Then the class of *primitive recursive* functions consists of just those functions which can be generated

by substitution according to i) above
by recursion according to ii) above
from the successor function **S**
from constant functions **C** (such that $\mathbf{C}(x) = 0$ for all non-negative x)
and from identity functions **U** (such that $\mathbf{U}_{in}(x_1 \ldots x_n) = x_i$, where $i \leq n$)

In other words, ϕ is primitive recursive if there exists a finite sequence of functions $\phi_1 \ldots \phi_n$ which terminates with ϕ such that each function of the sequence is (with respect to preceding functions), as i) above, or as ii) above, or is itself **S**, or **C**, or **U**.

6. Recursive functions have this very important property: for each given set of values of the arguments, the value of the function can be computed by a finite procedure.

7. A *relation* is recursive if the function representing it is recursive. Hence recursive relations (i.e. classes) *are decidable*: for each set of natural numbers, it can be determined by a finite procedure whether or not the relation holds (i.e. whether or not the number belongs to the class).

'*General Recursive*': The discussion of primitive recursion above was only preparatory. It is general recursion (or its equivalent) which is required in any proof of the Gödel theorem.

1. The notion of primitive recursion rests on what is involved in showing a *single* function ϕ to be primitive recursive. But consider a situation in which *several* functions are introduced simultaneously by a single set of recursion equations (as if Peano had defined '+' and '×' simultaneously, in terms of a function ϕ, by a complex of three equations).

2. Suppose $\psi\,(y)$ and $\pi\,(x)$ are given (primitive) recursive functions. And suppose we seek to define the function $\phi\,(x, y)$ by the relations:

$$\phi\,(0, y) = \psi\,(y)\ ,$$

$$\phi\,(x+1, 0) = \pi\,(x)\ ,$$

$$\phi\,(x+1, y+1) = \phi\,(x, \phi\,(x+1, y))$$

[This is an example of definition by induction with respect to two variables simultaneously.] But, as so defined, the function ϕ is not *in general* recursive in the limited sense of 5, ii above. To obtain precision with respect to such functions ϕ we must generalize our idea of 'recursive'; we must make the scheme of 5 above more general. The consideration of various sorts of functions defined by inductions leads to this question: what would one mean by the expression '*every* recursive function'?

3. As a first approximation this definition may help:

if ϕ denotes an unknown function, and $\psi_1 \ldots \psi_n$ are known functions, and if the ψ's and the ϕ are substituted in, or for, one another (in the most general manner) and certain pairs of the resulting expressions are equated, then if the resulting set of functional equations has one and only one solution for ϕ, ϕ is a (general) recursive function.

Thus if in
$$\phi\,(x, 0) = \psi_1\,(x)$$
$$\phi\,(0, y\ 1) = \psi_2\,(y)$$
$$\phi\,(1, y\ 1) = \psi_3\,(y)$$
$$\phi\,(x\ 2, y\ 1) = \psi_4\,(\phi\,(x, y + 2), \phi\,(x, \phi\,(x, y + 2)))$$

there proves to be but one and only one solution for ϕ, then ϕ is general recursive.

4. More exactly, $\phi\,(x)$ is general recursive if

i) there are two primitive recursive functions π and ψ such that

$$(\perp y)\,(\exists x)\,(\psi\,(y, x) = 0), \quad \text{and}$$

ii) $$\phi\,(x) = \pi\,(\mu\,k, \psi\,(x, k) = 0)$$

Thus $\phi\,(x)$ is general recursive if there exists a set of equations from which it could be *mechanically* deduced that for

$$\phi\,(0) = k_0$$
$$\phi\,(1) = k_1$$[4]

'General Recursive' and 'Effective': This 'mechanical' feature of determining whether a function $\phi\,(x)$ is general recursive is very important. It begins to suggest how the ideas of 'effectiveness' and 'decision procedure' tie in with certain aspects of general recursive functions. We will anticipate the next two sections with a few general remarks here.

We require a decision procedure to be mechanical, – independent of an individual's ingenuity. Unless it provided an effective, mechanical way of computing the truth-value of all of its possible formulae the propositional calculus would not be said to possess a decision procedure, which (in the form of the truth table) it does. But we have just seen that a function $\phi\,(x)$ is general recursive only if it can be derived mechanically from an existing set of equations by the most general recursive methods. This immediately suggests the following possibility: A formal system will be described as 'complete' only if it fulfills certain requirements. It might be possible to express these requirements in terms of the properties of some suitably specified (numerical) function ϕ which, if ϕ can be shown to be general recursive this will be tantamount to demonstrating the completeness of the system. That is, the system will possess a ϕ capable of indicating whether (for any well-formed formula $\mathfrak{F}$) $\mathfrak{F}$ or $\sim \mathfrak{F}$ is a theorem. [If, however, ϕ is demonstrably *not* derivable from the equations of the system, the latter will have been shown to lack a decision procedure, and hence will be incomplete.]

With the aid of Gödel's ingenious device of representing formulae and sequences of formulae by means of numbers (to be explained in the next section), it is possible to define 'effectiveness' in the more exact sense suggested above. This is done by making the idea correspond to that of the *recursiveness of numerical functions.* Hence the decision problem of Arithmetic[5] is just the problem of determining a general recursive function of an argument such that if the argument of the function is the Gödel number of a (well-formed) formula $\mathfrak{F}$ then the value of the function must be either *0* (i.e. $\mathfrak{F}$ is a theorem or axiom) or *1* ($\mathfrak{F}$ is *not* a theorem or axiom, but $\sim \mathfrak{F}$ *is*), but not both of course. The traditional assumption had been that in any consistent formal mathematical system every well-constructed $\mathfrak{F}$ was decidable, i.e. either $\mathfrak{F}$ or $\sim \mathfrak{F}$ could be proved to be a theorem or an axiom. It was also assumed that for every such system there could be discovered a technique (like the truth-table procedure) by the use of which one could determine of any $\mathfrak{F}$, whether it or $\sim \mathfrak{F}$ was a theorem or axiom of the system. This was the assumption that every such system was complete.

Gödel proved that there is no general recursive function of the type specified above. If there *were*, Arithmetic would be inconsistent. There are Arithmetic formulae $\mathfrak{F}$ which, though well-formed and hence mathematically meaningful,[6] are yet not decidable; neither $\mathfrak{F}$ nor $\sim \mathfrak{F}$ can be demonstrated as a theorem. Hence can be no generally effective decision procedure for Arithmetic, – Arithmetic is incomplete. The next section will make possible an elucidation and proof of these important theses.

III. THE ARITHMETIZATION OF LOGICAL SYNTAX

Any proof of Gödel's Theorem requires the use of two logics. One of these is the logic of the formal system which includes Arithmetic. ('the Arithmetic system'). This is the system *about which* the Theorem is proved. The other logic is that *in which* the Theorem is proved.

One of Gödel's discoveries was of a single notation in which both these logics could be expressed. In what follows, the *meaning* of the symbols involved is immaterial. It is even desirable that questions of meaning should be ignored or suppressed. Notions which have to do with the *purely formal* aspects of the Arithmetic system may be called 'metamathematical'. In other words, we are going to undertake a metamathematical proof of a theorem (Gödel's) *about* the Arithmetic system. Gödel expresses the logics of both the Arithmetic and the meta-arithmetic systems in one notation, as follows: we require a logic consisting of these symbols:

0 'zero' (the number)

S '. . . is the successor of . . .'

= '. . . equals . . .'

$\sim$	'not . . .'
v	'. . . or . . .'
.	'. . . and . . .'
$\rightarrow$	'. . . implies . . .'
$\equiv$	'. . . implies and is implied by . . .'
$\perp$	'all . . . such that'
$\exists$	'there is at least one . . . such that'
μ	'the smallest . . . such that'
(	'bracket' (opening)
)	'bracket' (closing)

plus an infinite set of variables in each of an infinite set of types (e.g. for propositions, A, B, C . . . ; for numbers, x, y, z . . . ; for functions, π, ϕ, ψ, . . . ; for classes, Π, Φ, Ψ, . . . etc.)

Numbers are now ordered these symbols:

0	*1*
S	*2*
=	*3*
$\sim$	*4*
v	*5*
.	*6*
$\rightarrow$	*7*
$\equiv$	*8*
$\perp$	*9*
$\exists$	*10*
μ	*11*
(	*12*
)	*13*
P_i^{α}	primes greater than *13* to variables of type α

Next, numbers must be assigned to formulae, as follows: suppose x_1, x_2, . . . x_n to be the numbers of the symbols of formula $\mathfrak{F}$ (in the order in which they occur in $\mathfrak{F}$). And let P_1, P_2, P_3, . . . P_n be the first n primes (in order of increasing magnitude). Then the number assigned to $\mathfrak{F}$ will be

$$P_1^{x_1}, \text{ times } P_2^{x_2}, \text{ times } P_3^{x_3}, \ldots \text{ times } P_n^{x_n}$$

For example, let us arithmetize in this manner the formula

$$(x)\ (\exists\, y)\ (\mathbf{S}\ (y,\ x))$$

['Every number has a successor'] According to our numbering scheme above the symbols of this formula may be numbered thus:

(	x	)	(	$\exists$	y	)	(	**S**	(	y	x	)	)
12,	*289,*	*13,*	*12,*	*10,*	*17,*	*13,*	*12,*	*2,*	*12,*	*17,*	*289,*	*13,*	*13.*

so that the *number* of the formula itself is

2^{12} *times* 3^{289} *times* 5^{13} *times* 7^{12} *times* 11^{10} *times* 13^{17} *times* 17^{13} *times*

19^{12} *times* 23^{2} *times* 29^{12} *times* 31^{17} *times* 37^{289} *times* 41^{13} *times* 43^{13}

an enormous number, – but nonetheless uniquely computable as *the number* of the formula in question. Let us call this the Gödel number ('the **G**-number') of $(x)\ (\exists\, y)\ (\mathbf{S}\ (y,\ x))$. For every *possible formula* within this logic a Gödel number is thus assigned. (This could be proved) And once a Gödel number *is* assigned to a formula, the formula can always be found as follows:

Factor the number into its prime factors.
Then the number of *2*-s occurring in the factorization is the number of the first symbol of the formula, the number of *3*-s occurring is the number of the second symbol of the formula, the number of *5*-s occurring is the number of the third symbol of the formula . . . etc.

Thus if the **G**-number of $(x)\ (\exists\, y)\ (\mathbf{S}\ (y,\ x))$ is factored into its prime factors there would be twelve *2*'s in the factorization, hence *12* would be the number of the first symbol of the formula, – '('. There would be two hundred and eighty nine *3*'s, so *289* is the number of the formula's second symbol, – 'x', . . . and so on. Similarly to each proof $\mathfrak{P}$ we order the integer which corresponds to the sequence of the integers ordered to the member-formulae of $\mathfrak{P}$. Then a one-to-one correspondence is determined between formulae (proofs) and a subset of the positive integers. After **G**-numbers have been assigned thus to formulae and proofs, one may proceed to define various metamathematical classes and relations of positive integers, – in fact there is one corresponding to each class and relation of formulae. (This too could be proved).

x then is a **G**-number if there is a formula, or proof, to which x corresponds in the following manner:

if Φ is a property of formulae, we can find a property ϕ of numbers such that a given formula $\mathfrak{F}$ has the property Φ if and only if the **G**-number of $\mathfrak{F}$ (in this case, x) has the property ϕ.

If the property Φ of formulae is properly chosen we can express *within* the Arithmetic system the proposition 'the number x has the property ϕ'.[7] Furthermore, if x is the **G**-number of a formula $\mathfrak{F}$ of the Arithmetic system, then we are also expressing *in that system* a metamathematic statement *about a formula* of that system. All this can be put in for form of the following basic lemma:

Let 'ϕ (x)' be expressible in the Arithmetic system. Then, if that system is suitably specified, see 'Proof' IV, 3), it contains a formula $\mathfrak{F}$ with a number n such that $\mathfrak{F}$ expresses 'ϕ (n)'. And since (in this system) n has the property ϕ if and only if $\mathfrak{F}$ [8] has the property Φ, it follows that $\mathfrak{F}$ expresses ' $\mathfrak{F}$ has the property Φ'. This is an expression *in* the arithmetic system which expresses a statement *about* a formula *of* the system. With this lemma, and the technique of Arithmetization, we can now proceed to express *in the notation of* the Arithmetic system many metamathematical *statements about* the system.

For example, suppose I wish to assert that the Arithmetic system is free from contradiction. And suppose that ϕ is such that $x \ \phi \ y$ shall mean that x and y are **G**-numbers and that the proof of which x is the **G**-number is the proof of the formula of which y is the **G**-number. Then, as a proposition *of* the Arithmetic system 1 can write:

$$(x,\ y,\ z)\ [\sim (x \ \phi \ z \ \text{ and } \ y \ \phi' \ \text{Neg.}\ (z))]$$

i.e., for all natural numbers x, y, and z, it is not the case that x represents a proof of the formula $\mathfrak{F}$, while y represents a proof of $\sim \mathfrak{F}$; – z is the **G**-number of $\mathfrak{F}$. This formula, of course can be arithmetized still further (as was $(x)\ (\exists\ y)\ (\mathbf{S}\ (y,\ x))$ above). As stated, it has a uniquely computable **G**-number. We are now prepared to prove Gödel's Theorem, – and to this important task we turn directly.

IV. INFORMAL PROOF

1. Assume that there *is* a decision procedure in Arithmetic; – assume, that is, that any formal system which includes Arithmetic is complete. Then, every well-formed formula $\mathfrak{F}$ of the system is decidable. Every $\mathfrak{F}$ either is, or is not, a theorem, – and this is, in principle, demonstrable.
2. Then, according to our account of decision procedure (I) and recursive functions (II), there is a general recursive function ϕ such that for any (well-formed) formula $\mathfrak{F}$ of which n is the **G**-number, either

$$\left.\begin{array}{l} \phi\ (n) = 0 \ \ \text{(i.e. the expression of which } n \text{ is the} \\ \qquad\qquad\ \ \text{Gödel number } \textit{is} \text{ a theorem)} \\ \\ \quad \text{or} \\ \\ \phi\ (n) = 1 \ \ \text{(i.e. the expression of which } n \text{ is the} \\ \qquad\qquad\ \ \text{Gödel number } \textit{is not} \text{ a theorem} \end{array}\right\} \begin{array}{l} \text{i.e. every} \\ \text{well-formed} \\ \mathfrak{F} \text{ is decidable} \end{array}$$

but not both, of course, – not if Arithmetic is to be consistent. According to an important lemma of Gödel[9] all general recursive functions are expressible in the notation of the Arithmetic system. Thus ϕ (n) can be expressed in purely arithmetic notation.

3. It was stated earlier that the Arithmetic system must be 'suitably specified'. It must, that is, fulfill the following condition:

   $\psi\ (p,\ p) = q$ is expressible (is well-formed) in the system,[10] – where $\psi\ (x,\ y)$ is the function described as follows.

4. $\psi\ (x,\ y)$ is the **G**-number of the formula $\mathfrak{F}$ got by taking the formula of which x is the **G**-number and replacing in that formula every free occurrence of v with the number y.[11]

   Furthermore, the arithmetical analogue of the syntactical operation of substitution can be defined rigorously within the notation of the Arithmetic system, according to the program set out in section III.[12] So the operation of replacement (substitution) required in 4. is expressible in arithmetic.

5. Thus $\psi\ (x,\ y)$ is the **G**-number of a well-formed formula $\mathfrak{F}$[13] and may be dealt with as follows:

   replace x with p
   replace y with p

   Thus:

   $$\psi\ (x,\ y) \equiv \psi\ (p,\ p)$$

   Thus $\psi\ (p,\ p)$ is the **G**-number of $\mathfrak{F}$

6. But, by the same argument, – now replacing p with v:

   $$\psi\ (p,\ p) \equiv \psi\ (\mathsf{v},\ \mathsf{v})$$

   Thus $\psi\ (\mathsf{v},\ \mathsf{v})$ is the **G**-number of $\mathfrak{F}$

7. So $\psi\ (\mathsf{v},\ \mathsf{v})$ is a number (by 6)
   And ϕ is expressible in arithmetic notation (by 2)
   Then, if $\phi\ (n) = 1$ is a well-formed formula of arithmetic (which by 1 and 2 it surely is) then

   $\phi\ [\psi\ (\mathsf{v},\ \mathsf{v})] = 1$ is a well-formed formula of Arithmetic.[14]

8. But since $\phi\ [\psi\ (\mathsf{v},\ \mathsf{v})] = 1$ is an Arithmetic formula, it must have its own **G**-number (Cf. section III above), – and we shall say that this number is p.

   The **G**-number of $\phi\ [\psi\ (\mathsf{v},\ \mathsf{v})] = 1$ then, is p.

9. Again, v may be replaced by p, giving

   $$\phi\ [\psi\ (p,\ p)] = 1$$

   an Arithmetic formula whose **G**-number is q.

Recapitulation

$\phi\ [\psi\ (\mathsf{v},\ \mathsf{v})] = 1$ has p as its **G**-number

and

$\phi\ [\psi\ (p,\ p)] = 1$ has q as its **G**-number

So that if the formula of which q is the **G**-number *were a theorem* we could assert (by 2)

$$\phi\ (q) = 0$$

From this we could assert, *as a theorem*, the formula

$$\phi\ [\psi\ (p,\ p)] = 1$$

which is, after all, just the formula of which q is the **G**-number, – a formula we have just supposed to be a theorem. (These moves are important for steps 13 and 14 below).

10. Refer now to step 4: it described $\mathfrak{F}$ as the formula got by taking the formula of which p is the **G**-number (we now know this to be $\phi\ [\psi\ (\mathrm{v}, \mathrm{v})] = 1$) and replacing in that formula every free occurrence of v with the number p. The result of this substitution would be

$$\phi\ [\psi\ (p,\ p)] = 1$$

so *this* is the formula $\mathfrak{F}$.

11. But $\psi\ (x,\ y)$ is the **G**-number of $\mathfrak{F}$ (by 4)

Hence $\psi\ (p,\ p)$ is the **G**-number of $\mathfrak{F}$ (by 5)

However, q is the **G**-number of $\mathfrak{F}$ (by 9)

Therefore,

$$\psi\ (p,\ p) = q$$

(by the 'principle' that two numbers which are the **G**-number of the same formula are the same number, Cf. section III).

Recapitulation

(by 8)	$\phi\ [\psi\ (\mathrm{v},\ \mathrm{v})] = 1$	has p as its **G**-number
(by 9)	$\phi\ [\psi\ (p,\ p)] = 1$	has q as its **G**-number
(by 11)	$[\psi\ (p,\ p)]$	$= q$

12. Refer now to our assumptions in steps 1 and 2.

There is a decision procedure in Arithmetic; every well-formed $\mathfrak{F}$ is such that, if n is its **G**-number.

$$\left.\begin{array}{l}\text{either } \phi\ (n) = 0 \\ \text{or } \phi\ (n) = 1\end{array}\right\} \begin{array}{c}\text{but not both} \\ \text{(i.e. Arithmetic is consistent)}\end{array}$$

Well, $\phi\ [\psi\ (p,\ p)] = 1$ is a well-formed $\mathfrak{F}$ (by 7), and q is its **G**-number. So either $\phi\ (q) = 0$, or $\phi\ (q) = 1$.

13. Suppose that

$\phi\ (q) = 0$ (i.e. the $\mathfrak{F}$ of which q is the **G**-number *is* a theorem)

(hence by 9)

$\phi\ [\psi\ (p,\ p)] = 1$ (i.e. this is just the assertion *as a theorem* of the $\mathfrak{F}$ of which q is the **G**-number)

(but then substituting, by 11)

$\phi\ (q) = 1$ (the $\mathfrak{F}$ of which q is the **G**-number *is not* a theorem)

13, however, proves only that $\phi\ (q) = 1$, by the simple principle of *reductio ad absurdum*:

$$\text{if } (\phi\ (q) = 0) \text{ entails} \sim (\phi\ (q) = 0), \text{ then} \sim (\phi\ (q) = 0).$$

[i.e. [$\phi\ (q) = 1$]]

We may at this point regard $\phi\ (q) = 1$ *as a theorem*, a formula which we can prove even when we assume its contradictory (as was done in 13).

To establish the *incompleteness* of the Arithmetic system (i.e. that the system lacks a decision procedure) we must also prove:

$$[(\phi\ (q) = 1) \rightarrow \sim (\phi\ (q) = 1)] \rightarrow \sim (\phi\ (q) = 1),$$

[i.e. $\phi\ (q) = 0$]

This we now proceed to do, making the *reductio ad absurdum* argument of 13 'complex'.

14. (by 13)

$\phi\ (q) = 1$ (i.e. this is a theorem which states that the $\mathfrak{F}$ of which q is the **G**-number *is not* a theorem)

(hence by substituting, 11)

$\phi\ [\psi\ (p, p)] = 1$ (i.e. since $q = \psi\ (p, p)$ we have here the *same* theorem.) But since the **G**-number of this theorem is q (by 9)

it follows that

$$\phi\ (q) = 0$$

15. Thus if '$\phi\ (q) = 0$' is represented by '$\mathfrak{P}$', we have proved that

$$[(\mathfrak{P} \rightarrow \sim \mathfrak{P}) \rightarrow \sim \mathfrak{P}] \text{ and } [(\sim \mathfrak{P} \rightarrow \mathfrak{P}) \rightarrow \mathfrak{P}]$$

If there were a decision procedure in the Arithmetic system (if there were a ϕ such that when n is the **G**-number of a well-formed $\mathfrak{F}$, either $\phi\ (n) = 0$ or $\phi\ (n) = 1$) then Arithmetic would be radically inconsistent.

Hence if the Arithmetic system is consistent[15] *then it cannot contain a decision procedure. Arithmetic formulae* $\mathfrak{F}$ *can always be constructed such that neither* $\mathfrak{F}$ *nor* $\sim \mathfrak{F}$ *is a theorem.*

This is Gödel's Theorem and we may now take it as proved. A few general remarks about the importance of this discovery will conclude this exposition. A corollary of the Theorem is that in every formal arithmetical system a real number can be constructed which cannot be defined in the system. This is a proposition which few mathematicians before 1930 would have countenanced.[16] An enormously important consequence of the Theorem (which is implicit in steps 13-15 above, but will not be proved separately) is this: *the consistency of a formal mathematical system can never be demonstrated by the methods of that system.*

Historically, this proposition is significant as having destroyed the metamathematical program known as 'Formalism'. Following Hilbert, the Formalists denied that the meaning of mathematical theorems accrues to them in any essential way from the nature of logic, a thesis to which the names of Frege, Russell and Whitehead had been appended.[17] For the Formalists pure mathematics is simply the science of the formal structure of symbols. The 'meaning' of mathematical theorems consists in exhibiting the structure of the systems in which they are theorems. So Hilbert and his collaborators came to conceive of all of mathematics in the form of rigorously symbolized theorems, deduced from (partially) uninterpreted axioms. The validity of these deductions was thought to be guaranteed by a second science, – 'metamathematics', – whose subject-matter is mathematics proper, and whose aim is to demonstrate the self-consistency of mathematics (using only elementary and indubitable methods of arithmetic). Could metamathematics but achieve this, mathematics proper would ensure its own validity. It could be regarded as an internally rigorous formal system of completely indeterminate external reference. It would exhibit by the multiplicity and interconnexion of its own symbols the structure of all possible formal systems. It looked as if Hilbert's program might succeed. A general proof of the consistency of Arithmetic was obviously the first requirement. And, indeed, this began to seem thoroughly demonstrable. It appeared that the consistency of elementary number theory (arithmetic, including logic), could be proved by a metamathematical inquiry which used only a part of the assumptions of the theory.

As has been shown, Gödel, making these same assumptions,[18] demonstrated the *impossibility* of proving the consistency of any elementary number theory within the theory itself. Instead, any consistency-proof requires new methods which are not expressible in the system itself. In the history of mathematical thought Gödel's discovery must rank with the Pythagorean's proof of the incommensurability of the diagonal to the side of a square. Both discoveries demanded profound readjustments in the very conception of the nature of mathematics. One further remark will conclude this paper. Gödel's Theorem should not be interpreted as a proof that there are mathematical problems which are finally and definitely undecidable. It asserts only that the concept *decidable* always refers to a *definite* formal system. If $\mathfrak{F}$ is undecidable in one particular system, the possibility is always open of constructing a more comprehensive system. However, there is no single system in which *all* arithmetic formulae could be decided, or in which *all* arithmetical concepts could be defined.

> 'All mathematics can be formalized; but mathematics can *never* be exhausted in any single system, – it requires an infinite series of discourses which get progressively more comprehensive.'[19]

So Gödel's Theorem might be paraphrased: there is no *completely* formal mathematical system.

NOTES

1. Gödel, K., Über die formal unentschiedbare Satze der Principia Mathematica und verwandter Systeme I, – *Monatshefte für Mathematik und Physik*, vol. XXXVIII (1931).

2. Russell, B., and Whitehead, A. N., (1910, 1925).

3. For E. L. Post (1936) and A. M. Turing (1937) an effective method of solving a set of problems exists if one could build a machine which would solve any problem of the set without human intervention (other than the insertion of the question and (later) the reading of the answer).

4. This is equivalent to saying that every recursive function is an effectively computable function, i.e. a function which can always be derived *mechanically* from some existing set of equations. This was proved by Church (1936) . . [It is in fact possible (due to a method discovered by Frege, – 'Begriffsschrift' p. 60, – but claimed later by both Peirce and Dedekind) to *prove* the actual existence of any function ϕ which satisfies the conditions expressed by an *admissible* set of recursive equations. ϕ may then be defined as: the function ϕ such that the recursion equations with suitable quantifiers prefixed, hold. The reader may find this fact helpful in IV, Cf. Kalmar (1940).]

5. Or rather, of any logical system which includes Arithmetic.

6. That these $\mathfrak{F}$ are well-formed follows from this: they are formulae built up recursively ('step by step') from elementary arithmetical propositions which are obviously well-formed, – indeed, pre-eminently so, for they are *just* tautologies. If these are not well-formed, then there is no stable conception whatever of what it is for a proposition to *be* well-formed. The undecidable $\mathfrak{F}$ which Gödel constructs are, as it were, *recursively reducible* to these indubitable propositions.

7. 'ϕ' can denote 'is divisible by y', 'is a prime number', 'is the *nth* prime', 'is the *nth* member of the sequence of positive integers which x represents', – and many more complicated (recursively-defined, arithmetically-expressible) properties. E.g. Let 'Subst $(\mathfrak{A}\,{}^{b}_{c})$' denote the expression obtained from $\mathfrak{A}$ by substituting c for each free occurrence of b in $\mathfrak{A}$. Then $\psi\,(x, y)$ may be a number having the property ϕ such that $\psi\,(x, y)$ is the **G**-number of the formula $\mathfrak{F}$, where $\mathfrak{F}$ is the result of Subst $(\mathfrak{B}^{p}_{v})$, – y being the **G**-number of the variable p, x being the **G**-number of the formula $\mathfrak{B}$. A ϕ with exactly this force is actually used in our informal proof; it is explained more fully in IV. It is mentioned here only to indicate that it is a ϕ rigourously expressible in the Arithmetic notation above.]

8. I.e. the formula of which n is the **G**-number.

9. Hinted at but not proved in section III, and assumed without discussion here.

10. This is no great assumption since $\psi (p, p) = q$, (or its equivalent), could be expressed in any logical system which might plausibly be said to include Arithmetic. $2 + 2 = 4$ would be one possible specification of $\psi (p, p) = q$.

11. p (the **G**-number) need not be distinguished from p (the ordinary arithmetic number). They have the same numerical value, i.e. they are the *same* number, as was explained in III.

12. For details of this see Gödel (1931), defs. 1-31, pp. 182-184.

13. $\mathfrak{F}$ will not actually be formulated until step 10, when its relations with previous steps will have become clear.

14. Doubts as to the well-formedness of $\phi\,[\psi\,(\mathrm{v}, \mathrm{v})] = 1$ may be resolved when it is seen that this formula is of exactly the same logical form (though not of the same logical type) as $4 - 3 = 1$. Substitute '2' for the vs in the formula such that $\psi\,(\mathrm{v}, \mathrm{v}) = 2 + 2$; then we have only to interpret 'ϕ' as *minus* **S** (v) to get the simple result $4 - 3 = 1$ (which is arithmetically well-formed if any formula is).

15. Which no one would deny; indeed this has been an article of faith underlying the whole of the proof above.

16. Cf. Hilbert, D. (1899, 1904, 1918, 1925, 1928).

17. Hilbert formalizes logic and arithmetic together, without taking logic as prior. In itself this is not opposed to the Frege-Russell-Whitehead view, since the choice of primitive symbols in the formalization can be made in more than one way. But for Hilbert, many of the theorems of the system are *ideale Aussagen*, – mere formulae without meaning in themselves. Frege and Russell, however, would give a meaning (as propositions of logic) to all formulae of the system. This is the central difference between the Formalist and the Logistic schools.

18. Incidentally, Gödel's assumptions do not differ markedly from those of Russell who 'solved' paradoxes of the Epimenides' variety by the restriction that no meaningful proposition can assert anything about itself. Gödel *does* construct propositions which assertions about themselves, but which, – since they are arithmetic formulae which involve only recursively defined functions, – are undoubtedly meaningful.

19. Carnap (1934).

BIBLIOGRAPHY

General

Carnap, R. — The Logical Syntax of Language, London (1934)

Frege, G. — *Begriffsschrift*, Halle (1879)

Frege, G. — *Die Grundlagen der Arithmetik*, Breslau (1884)

Frege, G. — *Grundgesetze der Arithmetik*, vol. I, Jena (1893), vol. II (1903)

Henkin, L. — *The Completeness of Formal Systems*, a dissertation, Princeton (1947)

Herbrand, J. — *Recherches sur la théorie de la démonstration*, Warsaw (1930)

Hilbert, D. — *Grundlagen der Geometrie*, Berlin (1930)

Hilbert, D. — *Über die Grundlagen der Logik und Arithmetik*, Verhaudlungen d. Math. Kongresses in Heidelberg (1904)

Hilbert, D. — *Axiomatisches Denken*, Mathematische Annalen (1918)

Hilbert, D. — *Über das Unendliche*, Mathematische Annalen (1925)

Hilbert, D. — *Die Grundlagen der Mathematik*, Abh. ans. d. Math. Sem. d. Hamb. Univ. (1928)

Hilbert, D. and Ackermann, W. — *Grundzüge der theoretischen Logik*, Berlin, (1928)

Hilbert, D. and Bernays, P. — *Grundlagen der Mathematik*, 2 vols., Berlin (1934, 1939)

Kalmar, L. — *On The Possibility of Definition By Recursion*, Acta Scientiarum Mathematicarum (1940)

Neumann, J. von. — *Eine Axiomatisierung der Mengenlehre*, Jour. für reine und angewandte Mathematik (1925)

Neumann, J. von. — *Zur Hilbertschen Beweistheorie*, Mathematische Zeitschrift (1927)

Peano, G. — *Formulaire de Mathematiques*, Turin, (1893-1908)

Presburger, M. — *Über die Vollstandigkeit eines gewissen Systems der Arithmetik . . .* etc. Comptes Rend. du I. Congrés des Math. des Pays Slaves (1929)

Post, El. — *Introduction to a General Theory of Elementary Propositions*, Amer. Jour. Math. (1921)

Post, El. — *Finite Combinatory Processes* – formulation 1, Journal of Symbolic Logic

Russell, B. and Whitehead, A. N. — *Principia Mathematica*, Vol. I, Cambridge, (1910)

Skolem, T. — *Über einige Satzfunktionen in der Arithmetik*, Skifter Mat. – Nat. Kl. (1930)

Tarski, A. — *A Decision Method for Elementary Algebra and Geometry*, Santa Monica (1948)

Formal Expositions of Gödel's Theorem

Carnap, R. — *Die Antinomien und die Unvollständigkeit der Mathematik*, Monatshefte für Mathematik und Physik (1934)

Church, A. — *An Unsolvable Problem of Elementary Number Theory*, Amer. Jour. Math. (1936)

Church, A. — *A Note on the Entschiedungsproblem*, Journal of Symbolic Logic (1936)

Chwistek, L. — *A Formal Proof of Gödel's Theorem*, Journal of Symbolic Logic (1939)

Gödel, K. — *Über die formal unentschiedbare Sätze der Principia Mathematica und verwandter Systeme I*, Monatshefte für Mathematik und Physik (1931) (also in translation, Princeton (1934))

Kleene, S. C. — *General Recursive Functions of Natural Numbers*, Mathematische Annalen (1936)

Kleene, S. C. — *Recursive Predicates and Quantifiers*, Trans. Amer. Math. Soc. (1943)

Rosser, B. and Kleene, S. C. — *On Undecidable Propositions of Formal Mathematical Systems* (Unpublished notes on lectures by Kurt Gödel) Princeton, (1934)

Turing, A. M. — *On Computable Numbers With An Application To The Entscheidungsproblem*, Proc. Lond. Math. Soc. (1937)

Turing, A. M. — *Computability and λ-definability*, Journal of Symbolic Logic (1937)

Informal Expositions of Gödel's Theorem

Carnap, R. — *The Logical Syntax of Language*, London (1937)

Findlay, J. — *Goedelian Sentences: a Non-numerical Approach*, Mind (1942)

Quine, W. V. — *Mathematical Logic*, New York (1940)

Quine, W. V. — *Methods of Logic*, New York (1950)

Rosser, B. — *An Informal Exposition of Proofs of Gödel's Theorems and Church's Theorem*, Journal of Symbolic Logic (1939)

Weyl, H. — *Philosophy of Mathematics and Natural Science*, Princeton (1949)

Indiana University
Bloomington, Indiana