L&PS – Logic and Philosophy of Science
Vol. IX, No. 1, 2011, pp. 151-158

Almost von Neumann, Definitely Gödel: The Second Incompleteness Theorem's Early Story

Giambattista Formica
Dipartimento di Filosofia, Letteratura, Storia e Scienze Sociali [FLESS]
Università degli Studi di Bari Aldo Moro
e-mail: giambattistaformica@tiscali.it

ABSTRACT. The discovery of the second incompleteness theorem has as its background the meeting between John von Neumann and Kurt Gödel at a Conference in Königsberg in 1930. After Gödel's announcement of an early version of the first incompleteness theorem there, von Neumann had a private discussion with him and some weeks later wrote to his colleague that he had proven the second incompleteness theorem. Unfortunately, Gödel had already submitted for publication his famous 1931 paper where the proof of the theorem was sketched. Once von Neumann knew this, he decided to leave to Gödel the paternity of the great discovery. In the literature there is some confusion over von Neumann's discovery, since his proof has been lost, and the issue should be considered open. In this paper I formulate a conjecture on von Neumann's discovery by analysing some basic documents.[1]

KEYWORDS: second incompleteness theorem, von Neumann's discovery, Königsberg conference, Gödel's announcement, Goldbach-like statements.

[1] I am very grateful to Fondazione CEUR for having sponsored this research.

Published by L&PS – Logic and Philosophy of Science
http://www2.units.it/ ~episteme — ISSN: 1826-1043

1. Introduction

On September 7th, 1930 Kurt Gödel announced an early version of his first incompleteness theorem: in the formal system of classical mathematics there are true but unprovable propositions.[2] The announcement was made during the roundtable on the foundations of mathematics that closed the International Conference on the *Epistemology of the Exact Sciences* held in Königsberg from the 5th to the 7th of September. Two days before, John von Neumann had presented at the same conference an address on the guidelines of the Hilbert program.[3] Gödel's announcement was bypassed by all the mathematicians present there with the only exception of von Neumann, who had a private discussion with him.[4]

On November 20th, von Neumann wrote Gödel that he had achieved, using Gödel's methods, a remarkable result: "that the consistency of mathematics is unprovable",[5] i.e., the second incompleteness theorem. Unfortunately, von Neumann's proof has been lost: however, in the letter he sketched the argument employed. In contrast, Gödel's reply (which has been lost, too) contains the statement that he had discovered the same fact, too, and an article with both theorems—the first and second incompleteness theorems—was already submitted to a journal.[6]

On November 29th, von Neumann wrote again saying two things of great interest: *i*) since he reached the theorem using Gödel's methods he will leave him the paternity of the discovery; *ii*) knowing his colleague's argument he is able to say he has employed a different one. "You proved $W \rightarrow A$, I showed independently the unprovability of W, in fact with a different argument".[7]

With this letter the story of the discovery of the second incompleteness theorem ends, and another story starts about its interpretation, analysed by Sieg.[8] Even though there is not yet a complete story about von Neumann, Gödel and the second incompleteness theorem, many questions that such a story would need to explore have already been answered.

Only one question still remains unanswered. How did von Neumann reach the second theorem knowing very few things about the first one? An argument

[2] Cf. Gödel (1931a), p. 203.
[3] Cf. von Neumann (1930).
[4] Cf. Wang (1981), pp. 654–5.
[5] Von Neumann (1930a), p. 337.
[6] Viz., Gödel (1931).
[7] Von Neumann (1930b), p. 339.
[8] Cf. Sieg (2005).

has been advanced by Franzén[9] but it seems historically inaccurate since it can be refuted using the basic documents. In the present paper I try to address the question, advancing a new conjectural argument.

In another paper von Neumann's discovery will be further detailed from the historical and technical point of view.

2. The announcement by Gödel

The announcement by Gödel is strictly related to the instrumentalist reading of the Hilbert program offered by von Neumann at the conference. According to the latter, consistency is a sufficient condition for showing the usefulness of transfinite systems for classical mathematics because these systems (if consistent) will allow the derivation of true finitist statements provable also by finitist methods alone. These systems will then be nothing but conservative extensions of systems with finitist contents. In other words, consistency will guarantee that no false theorems of the (old) finite domain can be proved in the new transfinite one. Gödel's argument is the following:

> According to the formalist view one adjoins to the meaningful propositions of mathematics transfinite (pseudo-)assertions, which in themselves have no meaning, but serve only to round out the system […]. This view presupposes that if one adjoins to the system *S* of meaningful propositions the system *T* of transfinite propositions and axioms and then proves a theorem of *S* by making a detour through theorems of *T*, this theorem is also contentually correct, hence through the adjunction of the transfinite axioms no contentually false theorems become provable. This requirement is customarily replaced by that of consistency. Now I would like to point out that one cannot, without further ado, regard these two demands as equivalent. […]. [A]s soon as one interprets the notion "meaningful proposition" sufficiently narrowly (for example, as restricted to finitary numerical equations), […] it is quite possible, for example, that one could prove a statement of the form $(Ex)F(x)$, where *F* is a finitary property of the natural numbers (the negation of Goldbach's conjecture, for example, has this form), by the transfinite means of classical mathematics, and on the other hand could ascertain by means of contentual considerations that all numbers have the property not-*F*; indeed, and here is precisely my point, this would still be possible even if one had demonstrated the consistency of the formal system

[9] Cf. Franzén (2005), pp. 97–8.

> of classical mathematics. For of no formal system can one affirm with certainty that all contentual considerations are representable within it.[10]

Gödel's argument directly leads to the first incompleteness theorem. In the case of finitary numerical equations, he says, it is possible to formally prove, by the transfinite means of a system *S* for classical mathematics, a false statement of the form $\exists x F(x)$, where *F* is a computable property of the natural numbers and $\exists x F(x)$ is the negation of a true Goldbach-like Statement (GlS) with the logical form $\forall x \neg F(x)$.

Goldbach-like Statements have particular properties. If a GlS $\forall x \neg F(x)$ is false, it can be contentually disproved by a computation that finds the first counterexample $F(a)$. And if the computation can be represented in *S*, then one will also have a formal proof of $\exists x F(x)$. The analogous statement cannot be said if a GlS is true.[11] It might be possible that a GlS is true—i.e., that it can be contentually verified that $\neg F(a)$, $\neg F(b)$, $\neg F(c)$, … holds for each *n*—but cannot be proved in *S*.

Gödel's argument takes into account exactly these properties. It is possible to have in *S* a formal proof of $\exists x F(x)$—i.e., the negation of a GlS $\forall x \neg F(x)$—and at the same time for it to be the case that for any given *n*, if one performed the appropriate computation for that *n*, to check (contentually) that $\neg F(x)$ holds for that *n*. Such a situation could hold even if one has proved that *S* is consistent. In fact, *S* remains syntactically consistent even if it proves the false proposition $\exists x F(x)$ but cannot prove its (true) negation $\forall x \neg F(x)$.

The last fact has a direct consequence. If *S* is consistent or there is a consistency proof for *S*, then there can be exhibited Goldbach-like Statements which are true but not provable in *S*. Then, we have two possibilities: either *S* proves false theorems against the instrumentalist reading of the Hilbert program or *S* is incomplete, e.g., he can neither prove the (contentually) true GlS $\forall x \neg F(x)$ nor its (contentually) false negation $\exists x F(x)$. I shall only note that these considerations describe the strategy Gödel employed in his 1931 paper to prove the first incompleteness theorem.[12]

The announcement by Gödel then states: "(Assuming the consistency of classical mathematics) one can even give examples of propositions (and in fact of those of the type of Goldbach or Fermat) that, while contentually true,

[10] Gödel (1931a), p. 201.

[11] For properties of Goldbach-like Statements and their relation to Gödel's theorems see Franzén (2005): in part., ch. 2, §§ 2.1–2.2.

[12] Gödel (1931), pp. 176–9.

are unprovable in the formal system of classical mathematics. Therefore, if one adjoins the negation of such a proposition to the axioms of classical mathematics, one obtains a consistent system in which a contentually false proposition is provable".[13]

Now I shall pay attention to the letter von Neumann wrote Gödel on November 20th. I would like to show that if one interprets this letter in light of Gödel's argument, one gets a clear idea of von Neumann's discovery.

3. The new argument

Let me quote the argument sketched by von Neumann in his letter of November 20th:

> In a formal system that contains arithmetic it is possible to express, following your considerations, that the formula 1 = 2 cannot be the end-formula of a proof starting with the axioms of this system—in fact, this formulation is a formula of the formal system under consideration. Let it be called *W*.
>
> In a contradictory system any formula is provable, thus also *W*. If the consistency [of the system] is established intuitionistically, then it is possible, through a "translation" of the contentual intuitionistic considerations into the formal [system], to prove *W* also. [...]. Thus with unprovable *W* the system is consistent, but the consistency is unprovable.
>
> I showed now: *W* is always unprovable in a consistent system, i.e., a putative effective proof of *W* could certainly be transformed into a contradiction.[14]

Its starting point affirms that in a formal system *S* that contains arithmetic it is possible to construct a formula *W* which expresses the metamathematical claim "*S* is consistent" ($Cons(S)$). In particular, *W* represents in *S* that "1 = 2 cannot be the end-formula of a proof starting from the axioms of the system".

In *S*, the logical form of *W* will then be $\forall x \neg F(x)$. If anyone doubts that *W* can have this form, one can see it expressing "for every *x*, *x* is not the Gödel-number of a proof of $1 = 2$ in *S*":

[13] Gödel (1931a), p. 203.
[14] Von Neumann (1930a), p. 337.

$$\left.\begin{array}{l}\ldots\\ \ldots\\ \ldots\\ \ldots\\ \ldots\\ \ldots\\ 1=2\end{array}\right\}\ (\mathrm{x})$$

Note that if one interprets $\forall x \neg F(x)$ in this way, *W* is a Goldbach-like Statement. In fact, $\neg F$ is a computable property because being the Gödel-number of a proof in *S* is so and *x* ranges over the natural numbers.[15]

(GlS) $Cons\ (S) :\ W \equiv\ \forall x \neg F(x)$

Now, let us see what happens if I apply to this fact, namely that *W* is a GlS, the considerations Gödel develops in his announcement. I shall briefly analyse these considerations under the three particular cases mentioned by von Neumann in his letter: 1) that *S* is inconsistent; 2) that *S* is consistent; 3) that there is a consistency proof for *S*.

1) Assumption: $\neg Cons(S)$. Thus $Cons(S)$ is false. If a GlS is false, it can be disproved by a computation, and if the computation can be represented in *S*, then it can be formally proved in *S* that $\neg\ Cons(S)$, namely $\exists x F(x)$. Furthermore, if *S* is inconsistent, one can formally prove any formula, thus also (*W*) $\forall x \neg F(x)$ which expresses $Cons(S)$. But, of course, one does not want *S* to be inconsistent. Thus, for this one needs first to assume that $Cons(S)$ is unprovable in *S*. In fact, if $Cons(S)$ is true and unprovable in *S*, then *S* is consistent.

2) Assumption: $Cons(S)$. As just said, *S* is consistent if $Cons(S)$, namely (*W*) $\forall x \neg F(x)$, is true and unprovable in *S*. However, this alone is not enough. In fact, for *S* to be consistent, $\exists x F(x)$ also has to be unprovable in *S*, since the latter expresses $\neg Cons(S)$. Therefore, *S* is consistent if neither $\forall x \neg F(x)$ nor $\exists x F(x)$ are provable in *S*, i.e., if $Cons(S)$ is true and undecidable in *S*.

[15] Cf. Gödel (1931), p. 171.

3) Assumption: $Pr(Cons(S))$, i.e., $Cons(S)$ is provable. By assumption there is in *S* a formal derivation of $Cons(S)$. Note that the derivation by itself does not imply that *S* is consistent because—as shown in 1)—$Cons(S)$ can be proved in *S*, also if *S* is inconsistent. Thus, only the unprovability of $Cons(S)$, i.e., (*W*) $\forall x \neg F(x)$, guarantees that *S* is consistent and one can conclude that *S* is consistent *if and only if* $Cons(S)$ is unprovable in *S*. From this conclusion the second incompleteness theorem can soon be obtained: if *S* is consistent, the consistency of *S* is unprovable in *S*, i.e., $Cons(S) \to \neg\, Pr(Cons(S))$. Hence, $Cons(S)$ is a true and unprovable GlS. Now, according to Gödel's announcement, given a true GlS $\forall x \neg F(x)$ and a provably consistent system *S*, it would be possible in *S* to prove $\exists x F(x)$, namely $\neg\, Cons(S)$, and not to prove $\forall x \neg F(x)$. Note that both facts would then contradict the assumption.

4. Concluding remarks

I would like to say now that the cases considered fit von Neumann's argument very well. In his letter he begins by saying: "In a formal system that contains arithmetic it is possible to express […] that the formula $1 = 2$ cannot be the end-formula of a proof starting with the axioms of this system—in fact, this formulation is a formula of the formal system […]. Let it be called *W*". This is just the starting point of my analysis. But then he adds: "In a contradictory system any formula is provable, thus also *W*. If the consistency [of the system] is established intuitionistically, then it is possible, through a 'translation' of the contentual intuitionistic considerations into the formal [system], to prove *W* also". This is exactly what I have explored in the first case. There one arrives to the conclusion: "[W]ith unprovable *W* the system is consistent, but the consistency is unprovable". In the end von Neumann adds one more note, that I have explored in the third case: "*W* is always unprovable in a consistent system, i.e., a putative effective proof of *W* could certainly be transformed into a contradiction".

In short my conjecture and conclusion is: von Neumann took very seriously the announcement by Gödel and having recognized $Cons(S)$ as a Goldbach-like Statement, he reached the second incompleteness theorem.

REFERENCES

FRANZÉN, T. (2005): *Gödel's Theorem. An Incomplete Guide to Its Use and Abuse*, Natick (MA): A.K. Peters.

GÖDEL, K. (1931): "On Formally Undecidable Propositions of *Principia Mathematica* and Related Systems I", in GÖDEL, K. (1986): *Collected Works. Volume I: Publications 1929–1936*, Oxford: Oxford UP, pp. 144–95.

GÖDEL, K. (1931a): "Discussion on Providing a Foundation for Mathematics", in GÖDEL, K. (1986): *Collected Works. Volume I: Publications 1929–1936*, Oxford: Oxford UP, pp. 200–5.

SIEG, W. (2005): "Only Two Letters: The Correspondence Between Herbrand and Gödel", *The Bulletin of Symbolic Logic*, 11, pp. 172–84.

VON NEUMANN, J. (1930): "The Formalist Foundations of Mathematics", in BENACERRAF, P. and PUTNAM, H. (eds.). (1964): *Philosophy of Mathematics. Selected Readings*, Englewood Cliffs (NJ): Prentice-Hall, pp. 50–4.

VON NEUMANN, J. (1930a): "Von Neumann to Gödel. Berlin, 20 November 1930", in GÖDEL, K. (2003): *Collected Works*. *Volume V: Correspondence H–Z*, Oxford: Oxford UP, pp. 336–9.

VON NEUMANN, J. (1930b): "Von Neumann to Gödel. Berlin, 29 November 1930", in GÖDEL, K. (2003): *Collected Works*. *Volume V: Correspondence H–Z*, Oxford: Oxford UP, pp. 338–41.

WANG, H. (1981): "Some facts about Kurt Gödel", *The Journal of Symbolic Logic*, 46, pp. 653–9.