12 The First Incompleteness Theorem, syntactic version

We now use the same construction of a Gödel sentence as in the previous chapter to show again that PA is incomplete, but this time making only syntactic assumptions. And then we show how to generalize this syntactic version of the incompleteness theorem.

12.1 ω-completeness, ω-consistency

We need to define two key notions. We'll assume in this section that we are dealing with theories whose language includes the language of basic arithmetic. And take all the quantifiers mentioned to run over the natural numbers.[1]

First, then,

Defn. 49. *A theory T is* ω-incomplete *iff, for some open wff $\varphi(\mathsf{x})$, T can prove $\varphi(\overline{\mathsf{n}})$ for each natural number n, but T can't go on to prove $\forall\mathsf{x}\varphi(\mathsf{x})$.*

We saw in §5.6 that Q is ω-incomplete: that's because it can prove each instance of $0 + \overline{\mathsf{n}} = \overline{\mathsf{n}}$, but can't prove $\forall\mathsf{x}(0 + \mathsf{x} = \mathsf{x})$. We added induction to Q hoping to repair as much ω-incompleteness as we could: but, as we'll see, PA remains ω-incomplete, assuming it is consistent.[2]

Second, we want the following idea:

Defn. 50. *A theory T is* ω-inconsistent *iff, for some open wff $\varphi(\mathsf{x})$, T can prove each $\varphi(\overline{\mathsf{n}})$ and T can also prove $\neg\forall\mathsf{x}\varphi(\mathsf{x})$.*

Or, entirely equivalently of course, we could say that T is ω-inconsistent if, for some open wff $\psi(\mathsf{x})$, $T \vdash \exists\mathsf{x}\psi(\mathsf{x})$, yet for each number n we have $T \vdash \neg\psi(\overline{\mathsf{n}})$.

Note that ω-inconsistency, like ordinary inconsistency, is a syntactically defined property: it is characterized in terms of what wffs can be proved by the theory, not in terms of what they mean. Note too that, in a classical context,

[1] If necessary, therefore, read $\forall\mathsf{x}\varphi(\mathsf{x})$ as a restricted quantifier $\forall\mathsf{x}(\mathsf{N}\mathsf{x} \to \varphi(\mathsf{x}))$, where 'N' picks out the numbers from the domain of the theory's native quantifiers (see Defn. 11).

[2] Why the 'ω' in 'ω-incomplete'? Because 'ω' is a standard label for the set of natural numbers (when equipped with their usual ordering).

ω-consistency – defined of course as not being ω-inconsistent! – trivially implies plain consistency. That's because T's being ω-consistent is a matter of its *not* being able to prove a certain combination of wffs, which entails that T can't prove *all* wffs, hence T can't be inconsistent.

Now compare and contrast. Suppose T can prove $\varphi(\overline{\mathsf{n}})$ for each n. T is ω-incomplete if it can't prove something we'd then also like it to prove, namely $\forall \mathsf{x}\varphi(\mathsf{x})$. While T is ω-inconsistent if it can actually prove the *negation* of what we'd like it to prove, i.e. it can prove $\neg\forall \mathsf{x}\varphi(\mathsf{x})$.

So ω-incompleteness in a theory of arithmetic is a regrettable weakness. But ω-inconsistency is a Very Bad Thing (not as bad as outright inconsistency, maybe, but still bad enough). For evidently, a theory T that can prove each of $\varphi(\overline{\mathsf{n}})$ and yet also prove $\neg\forall \mathsf{x}\varphi(\mathsf{x})$ is just not going to be an acceptable candidate for regimenting arithmetic.

Bring semantic ideas back into play for a moment. Suppose T's standard numerals denote the numbers and the quantifier here runs over the natural numbers. Then it can't be the case that each of $\varphi(\overline{\mathsf{n}})$ is true and yet $\neg\forall \mathsf{x}\varphi(\mathsf{x})$ is true too. So our ω-inconsistent T can't be sound.

Given that we want formal arithmetics to have theorems which *are* all true on a standard interpretation, we must therefore want ω-consistent arithmetics. And given that we think e.g. PA *is* sound on its standard interpretation, we are committed to thinking that it *is* ω-consistent.

12.2 The First Theorem for PA – the syntactic version

Remember Defn. 46: the wff Gdl by hypothesis doesn't just express *Gdl* but *captures* it. Using this fact about Gdl, we can again show that PA does not prove G, but this time *without* making the semantic assumption that PA is sound:

Theorem 40. *If* PA *is consistent,* PA $\nvdash$ G*.*

Proof. Recall, G is the diagonalization of the open wff $\mathsf{U} =_{\text{def}} \forall\mathsf{x}\neg\mathsf{Gdl}(\mathsf{x},\mathsf{y})$, i.e. is the wff $\forall\mathsf{x}\neg\mathsf{Gdl}(\mathsf{x},\overline{\ulcorner\mathsf{U}\urcorner})$ (see §11.2).

Suppose now that G *is* provable in PA. If G has a proof, then there is some super g.n. m that codes its proof. But G is the diagonalization of the wff with g.n. $\ulcorner\mathsf{U}\urcorner$. Hence, by definition, $Gdl(m,\ulcorner\mathsf{U}\urcorner)$.

Now we use the fact that Gdl captures the relation *Gdl*. Since $Gdl(m,\ulcorner\mathsf{U}\urcorner)$, by the definition of capturing we have (i) $\mathsf{PA}\vdash\mathsf{Gdl}(\overline{\mathsf{m}},\overline{\ulcorner\mathsf{U}\urcorner})$.

But since G is none other than $\forall\mathsf{x}\neg\mathsf{Gdl}(\mathsf{x},\overline{\ulcorner\mathsf{U}\urcorner})$, the assumption that G is provable is this: $\mathsf{PA}\vdash\forall\mathsf{x}\neg\mathsf{Gdl}(\mathsf{x},\overline{\ulcorner\mathsf{U}\urcorner})$. The universal quantification here entails any instance. Hence in particular (ii) $\mathsf{PA}\vdash\neg\mathsf{Gdl}(\overline{\mathsf{m}},\overline{\ulcorner\mathsf{U}\urcorner})$.

So, combining (i) and (ii), the assumption that G is provable entails that PA is inconsistent. Therefore, if PA is consistent, there can be no PA proof of G. □

Here's an immediate corollary of that last theorem:

Theorem 41. *If* PA *is consistent, it is ω-incomplete.*

Proof. Assume PA's consistency. Then we've shown that $\mathsf{PA} \nvdash \mathsf{G}$, i.e.,

1. $\mathsf{PA} \nvdash \forall \mathsf{x} \neg \mathsf{Gdl}(\mathsf{x}, \overline{\ulcorner \mathsf{U} \urcorner})$.

Since G is unprovable, that means that no number is the super g.n. of a proof of G. That is to say, no number numbers a proof of the diagonalization of U. That is to say, for any particular m, it *isn't* the case that $Gdl(m, \ulcorner \mathsf{U} \urcorner)$. Hence, again by the fact that Gdl captures Gdl, we have

2. For each m, $\mathsf{PA} \vdash \neg \mathsf{Gdl}(\overline{\mathsf{m}}, \overline{\ulcorner \mathsf{U} \urcorner})$.

Putting $\varphi(\mathsf{x}) =_{\text{def}} \neg \mathsf{Gdl}(\mathsf{x}, \overline{\ulcorner \mathsf{U} \urcorner})$, the combination of (1) and (2) therefore shows that PA is ω-incomplete. □

We can now easily derive

Theorem 42. *If* PA *is consistent, then* $\mathsf{PA} + \neg \mathsf{G}$ *(the theory you get by adding* $\neg \mathsf{G}$ *as an additional axiom) is also consistent but is* ω*-inconsistent.*

Proof. If $\mathsf{PA} + \neg \mathsf{G}$ were inconsistent, then PA would prove G by reductio: but Theorem 40 shows that PA doesn't prove G if it is consistent. So $\mathsf{PA} + \neg \mathsf{G}$ is consistent.

Now, by the definition of $\neg \mathsf{G}$,

1. $\mathsf{PA} + \neg \mathsf{G} \vdash \exists \mathsf{x} \mathsf{Gdl}(\mathsf{x}, \overline{\ulcorner \mathsf{U} \urcorner})$.

And since the expanded theory includes PA, we will have as in the proof for Theorem 40

2. For each m, $\mathsf{PA} + \neg \mathsf{G} \vdash \neg \mathsf{Gdl}(\overline{\mathsf{m}}, \overline{\ulcorner \mathsf{U} \urcorner})$.

And (i) and (ii) together imply that $\mathsf{PA} + \neg \mathsf{G}$ is ω-inconsistent. □

We'll now show that PA also can't prove the *negation* of G, again without assuming PA's soundness:

Theorem 43. *If* PA *is* ω*-consistent,* $\mathsf{PA} \nvdash \neg \mathsf{G}$.

Proof. Suppose $\neg \mathsf{G}$ is provable in PA. Then the theory PA would be equivalent to $\mathsf{PA} + \neg \mathsf{G}$. Hence, by the previous theorem, assuming PA is consistent, it would be ω-inconsistent. Hence, assuming PA is ω-consistent and hence consistent, it can't prove $\neg \mathsf{G}$. □

Now recall that G is a Π_1 sentence. That observation put together with what we've shown in this section gives us the following portmanteau result:

Theorem 44. *If* PA *is consistent, then there is a* Π_1 *sentence* G *such that* $\mathsf{PA} \nvdash \mathsf{G}$*, and if* PA *is* ω*-consistent* $\mathsf{PA} \nvdash \neg \mathsf{G}$*. Hence, assuming* ω*-consistency and so consistency,* PA *is negation incomplete.*

12.3 Generalizing the proof

The proof for Theorem 44 evidently generalizes. Suppose T is a p.r. axiomatized theory which (perhaps after introducing some new vocabulary by definitions) *contains* Q – in the obvious sense that the language of T includes the language of basic arithmetic, and T can prove every Q-theorem. Then, assuming we are working with normal scheme for Gödel-numbering wffs of T, the relation $Gdl_T(m, n)$ which holds when m numbers a T-proof of the diagonalization of the wff with number n will be primitive recursive again.

Since T can prove everything Q proves, T will be able to capture the p.r. relation Gld_T by a Σ_1 wff Gld_T. Just as we did for PA, we'll be able to construct the corresponding Π_1 wff G_T. And, exactly the same arguments as before will then show, more generally,

Theorem 45. *If T is a consistent p.r. axiomatized theory which contains Q, then there will be a Π_1 sentence G_T such that $T \nvdash \mathsf{G}_T$, and if T is ω-consistent, $T \nvdash \neg\mathsf{G}_T$. Hence, assuming ω-consistency and so consistency, T is negation incomplete.*

And note, this gives us another incompletability theorem: if we keep chucking more and more additional axioms at our theory T, it will still remain negation incomplete, unless it stops ω-consistent or stops being p.r. axiomatized.

When people refer to the *First Incompleteness Theorem* (without qualification), they typically mean something like our last Theorem, deriving incompleteness from *syntactic* assumptions. Let's re-emphasize that last point. Being p.r. axiomatized is a syntactic property; containing Q is a matter of Q's axioms being adopted or being derivable, a syntactic property; being consistent here is a matter of no contradictory pair φ, $\neg\varphi$ being derivable, being ω-consistent is another syntactic property as we stressed before. The chains of argument that lead to this theorem depend just on the given syntactic assumptions, via e.g. the proof that Q can capture all p.r. functions – another claim about a syntactically definable property. That is why we are calling this the *syntactic* incompleteness theorem.

Of course, we are *interested* in these various syntactically definable properties because of their semantic relevance: for example, we care about the idea of capturing p.r. functions because we are interested in what an interpreted theory might be able to prove in the sense of establish-as-true. But it is one thing for us to have a semantic motivation for being interested in a certain concept, it is another thing for that concept to have semantic content.

12.4 Comparing old and new syntactic incompleteness theorems

Compare Theorem 45 with our initially announced

Theorem 2. *Suppose T is a formal axiomatized theory whose language contains the language of basic arithmetic. Then, if T is consistent and can prove a certain* *modest amount of arithmetic (and has a certain additional property that any sensible formalized arithmetic will share), there will be a sentence* G_T *of basic arithmetic such that* $T \nvdash \mathsf{G}_T$ *and* $T \nvdash \neg\mathsf{G}_T$*, so* T *is negation incomplete.*

Our new theorem fills out the old one in various respects. It fixes the 'modest amount of arithmetic' that T is assumed to contain and it also spells out the 'additional desirable property' of ω-consistency which we previously left mysterious. Further it tells us more about the undecidable Gödel sentence – namely it has minimal quantifier complexity, i.e. it is a Π_1 sentence of arithmetic. Our new theorem is weaker, however, as it only applies to p.r. axiomatized theories, not to formal axiomatized theories more generally. But we've already noted, that that's not much loss. (And again, if we insist, we can in fact go on to make up the shortfall: see the next Interlude.)

12.5 Gödel's own Theorem

As we said, Theorem 45, or something like it, is what people usually mean when they speak without qualification of 'The First Incompleteness Theorem'. But since the stated theorem refers to Robinson Arithmetic Q (developed by Robinson in 1950), and Gödel didn't originally know about that (in 1931), our version can't be quite what Gödel originally proved. But it is a near miss. Let's explain.

Looking again at our analysis of the syntactic argument for incompleteness, we see that we are interested in theories which extend Q *because we are interested in theories which can capture p.r. relations like Gdl*. It's being able to capture *Gdl* that is the crucial condition for a theory's being incomplete. So let's say

Defn. 51. *A theory* T *is* p.r. adequate *if it can capture all primitive recursive functions and relations.*

Then, instead of mentioning Q, let's instead explicitly write in the requirement of p.r. adequacy. So, by just the same arguments,

Theorem 46. *If* T *is a p.r. adequate, p.r. axiomatized theory whose language includes* L_A*, then there is* Π_1 *sentence* φ *such that, if* T *is consistent then* $T \nvdash \varphi$*, and if* T *is* ω*-consistent then* $T \nvdash \neg\varphi$*.*

And *this* is pretty much Gödel's own general version of the incompleteness result. I suppose that it has as much historical right as any to be called *Gödel's* First Theorem.

('Hold on! If *that's* the original First Theorem, we didn't need to do all the hard work showing that Q and PA are p.r. adequate, did we?' Well, yes and no. No, proving *this* original version of the Theorem of course doesn't depend on proving that any particular theory is p.r. adequate. But yes, showing that this Theorem has real bite, showing that it does actually apply to familiar arithmetics, does depend on proving that these arithmetics are indeed p.r. adequate.)

Thus, in his 1931 paper, Gödel first proves his Theorem VI which, with a bit of help from his Theorem VIII shows that the formal system P – which is his simplified version of the hierarchical type-theory of *Principia Mathematica* – has a formally undecidable Π_1 sentence.[3] Then he immediately generalizes:

> In the proof of Theorem VI no properties of the system P were used besides the following:
>
> 1. The class of axioms and the rules of inference (that is, the relation 'immediate consequence') are [primitive] recursively definable (as soon as we replace the primitive signs in some way by the natural numbers).
> 2. Every [primitive] recursive relation is definable [i.e. is 'capturable'] in the system P.
>
> Therefore, in every formal system that satisfies the assumptions 1 and 2 and is ω-consistent, there are undecidable propositions of the form $[\forall x F(x)]$, where F is a [primitive] recursively defined property of natural numbers, and likewise in every extension of such a system by a recursively definable ω-consistent class of axioms.

Which gives us our Theorem 46.

[3] Or as Gödel put it, the undecidable sentence is 'of Goldbach type'. The allusion here is to Goldbach's conjecture that every even number other than two is the sum of two primes. The claim that the particular number n is an even number other than two and is the sum of two primes is expressible by a Δ_0 wff (why?). So Goldbach's conjecture, the universal quantification of this claim about n, is expressible by a Π_1 wff.