Gödel's 2nd Incompleteness Theorem

We outline how Gödel's second incompleteness theorem is formulated and proved. Essentially this theorem states that any "strong enough" axiomatizable theory can not prove that it is consistent. (Unless its inconsistent, in which case everything is provable).

Notation: Let A be a recursive set of axioms.

(a) $\mathrm{Deduct}_A(w, x)$ is a numeralwise determined formula of Q and represents the recursive relation:

"x is the Gödel number of a formula of A and w codes a deduction from A of the formula coded by x".

(b) $Thm_A(x)$ is the formula $\exists w(Deduct_A(w,x))$

So $Thm_A(x)$ asserts "x codes a consequence of A".

(c) $Consis(A)$ is the sentence

$$\neg \exists w(Deduct_A(w, S^{\#0=1} 0))$$

So $Consis(A)$ asserts "$A \nvdash 0=1$". This implies that A is consistent. If $Cn\,A \supseteq Q$ then $Consis(A)$ is true iff A is consistent.

<u>Definition:</u> (Peano arithmetic).

PA is a set of axioms in the language $0, S, +, \cdot, <$.

PA contains following sentences:

(a) The sentences in Q,

(b) For each formula φ, the <u>induction axiom</u>,

$$\varphi_0^{v_1} \wedge \forall v_1(\varphi \to \varphi_{Sv_1}^{v_1}) \to \forall v_1 \varphi$$

[Fact: $Cn\,PA$ is not finitely axiomatized.]

Actually, the choice of Peano arithmetic (PA) is not too crucial. The important properties of PA are:

① PA is consistent

② $Cn\ PA \supseteq Q$

③ PA satisfies propositions B-1 and B-2 below.

Thus the results below also hold for theories such at ZF (Set Theory). However, with Set Theory, some techical difficulties arise in interpreting formulas in the language of Q in the language of Set Theory - see Enderton for more details.

<u>Proposition B-1</u>: Let φ, ψ be arbitrary formulas.

Then

$$PA \vdash Thm_{PA}(S^{\#\varphi}0) \wedge Thm_{PA}(S^{\#(\varphi\to\psi)}0) \to$$
$$\to Thm_{PA}(S^{\#\psi}0).$$

In words, PA proves that modus ponens allows a deduction of φ and a deduction of $\varphi\to\psi$. to give a deduction ψ

<u>Proof</u>: A detailed proof of Proposition B-1 would be too long for us to carry out in full – So we shall just say why its reasonable.

Note that if

$$\mathrm{Deduct}_{PA}(v_0, S^{\#\varphi}0)$$

and $\mathrm{Deduct}_{PA}(w_0, S^{\#(\varphi\to\psi)}0)$

hold (with $v_0, w_0 \in \mathbb{N}$), then

$$\mathrm{Deduct}_{PA}(v_0 * w_0 * \langle \#\psi \rangle, S^{\#\psi}0)$$

is also true.

So it seems reasonable that

$$PA \vdash \mathrm{Deduct}_{PA}(v, S^{\#\varphi}0) \wedge \mathrm{Deduct}_{PA}(w, S^{\#(\varphi\to\psi)}0)$$

$$\to \exists x\, \mathrm{Deduct}_{PA}(x, S^{\#\psi}0).$$

since $x = v * w * \langle \#\psi \rangle$ will work.

<u>Proposition B-2</u>: Let φ be an arbitrary formula.

Then

$$PA \vdash Thm_{PA}(S^{\#\varphi}0) \rightarrow Thm_{PA}(S^{\#Thm_{PA}(S^{\#\varphi}0)}0)$$

<u>Proof</u>: Omitted - Lets show instead that

$$\mathfrak{N} \models Thm_{PA}(S^{\#\varphi}0) \rightarrow Thm_{PA}(S^{\#Thm_{PA}(S^{\#\varphi}0)}0)$$

So suppose $Thm_{PA}(S^{\#\varphi}0)$ is true in $\mathfrak{N}$.

Hence, $\exists w_0 \in N$ such that $\mathfrak{N} \models Deduct_{PA}(S^{w_0}0, S^{\#\varphi}0)$

Since $Deduct_{PA}$ is numeralwise determined and $Q \subseteq Th\,\mathfrak{N}$,

$$Q \vdash Deduct_{PA}(S^{w_0}0, S^{\#\varphi}0)$$

$\therefore$ So $Q \vdash Thm_{PA}(S^{\#\varphi}0)$. by existential rule

$\therefore$ Hence $PA \vdash Thm_{PA}(S^{\#\varphi}0)$ since $Q \subseteq Cn\,PA$

In other words, $\mathfrak{N} \models Thm_{PA}(S^{\#Thm_{PA}(S^{\#\varphi}0)})$.

Corollary B-3:

(a) If $PA \vdash \varphi$ then $Q \vdash Thm_{PA}(S^{\#\varphi}0)$

(b) If $PA \vdash \varphi$ then $PA \vdash Thm_{PA}(S^{\#\varphi}0)$.

Proof: This shown by the argument in support of B-2 on the previous page. (Note the lines marked with $\therefore$).

QED Corollary B-3

Definition: Let σ be a sentence such that

$$Q \vdash \sigma \leftrightarrow \neg Thm_{PA}(S^{\#\sigma}0).$$

σ exists by Gödel's Fixed Point Lemma applied to

$$\beta(v_1) = \neg Thm_{PA}(v_1).$$

Note σ says "I am not a consequence of PA"

or "I am not a PA-theorem".

<u>Proposition B-4</u>: σ is true.

That is to say, $\mathfrak{N} \models \sigma$.

<u>Proof</u>: Suppose σ is false. I.e., $\mathfrak{N} \models Thm_{PA}(S^{\#\sigma}0)$

In other words, $PA \vdash \sigma$.

So $PA \vdash \neg Thm_{PA}(S^{\#\sigma}0)$ (by defn of σ.)

On the other hand,

$PA \vdash Thm_{PA}(S^{\#\sigma}0)$ by Corollary B-3(b).

And this contradicts the consistency of PA.

<u>QED</u> Proposition B-4.

<u>Corollary B-5</u>: $PA \nvdash \sigma$

<u>Pf</u>: Since σ is true. <u>q.e.d</u>

<u>Proposition B-6</u>: $PA \vdash (\neg\sigma \rightarrow Thm_{PA}(S^{\#\sigma}0))$

<u>Pf</u>: By the definition of σ and since $Q \subseteq Cn\, PA$.

<u>Proposition B-7:</u> $PA \vdash (\neg\sigma \to Thm_{PA}(S^{\#(\neg\sigma)}0))$

<u>Proof:</u>

$PA \vdash \neg\sigma \to Thm_{PA}(S^{\#\sigma}0)$ by Proposition B6.

$PA \vdash Thm_{PA}(S^{\#\sigma}0) \to Thm_{PA}(S^{\#Thm_{PA}(S^{\#\sigma}0)}0)$

by Proposition B-2.

$PA \vdash Thm_{PA}(S^{\#(Thm_{PA}(S^{\#\sigma}0)\to\neg\sigma)}0)$ by def'n of σ

and Corollary B-3 (b).

$PA \vdash Thm_{PA}(S^{\#Thm_{PA}(S^{\#\sigma}0)}0) \to Thm_{PA}(S^{\#(\neg\sigma)}0)$

by Proposition B-1 and modus ponens

$PA \vdash \neg\sigma \to Thm_{PA}(S^{\#(\neg\sigma)}0)$ by modus ponens applied twice.

<u>QED</u> Proposition B-7.

<u>Proposition B-8:</u> $PA \vdash \neg\sigma \to \neg Consis(PA)$.

<u>Proof:</u> First note $PA \vdash Thm_{PA}(S^{\#(\varphi\to\neg\varphi\to 0=1)}0)$.

Now use Propositions B-7, B-8 and two uses of Proposition B-1 to get $PA \vdash \neg\sigma \to Thm_{PA}(S^{\#0=1}0)$.

QED

<u>Corollary B-9</u>: $PA \nvdash Consis(PA)$

<u>Proof</u>: By Corollary B-5 and Proposition B-8.

<u>Q.E.D.</u>

<u>MetaTheorem</u>: Let Z be any axiomatized theory extending PA (with the set of integers, $0, S, +, \cdot, <$ all definable in Z). Then $Z \nvdash Consis(Z)$.