The 2nd Incompleteness Theorem

↳ informally: says any sufficiently strong axiom system Q for arithmetic cannot prove its own consistency.

Observation to get started: we showed, if Q consistent then $Q \nvdash \sigma_Q$, where σ_Q is Gödel sentence (this is the direction where we didn't need ω-consistency)

i.e. we proved the "sentence"
Q consistent $\Rightarrow \neg Prvbl_Q(\langle \underline{\sigma_Q} \rangle)$

- Sps we could write this statement formally somehow in L_{ar}.
- and also formalize its proof in Q
i.e. show $Q \vdash$ "Q is consistent" $\Rightarrow \neg Prvbl_Q(\langle \underline{\sigma_Q} \rangle)$

- Then if $Q \vdash$ "Q is consistent"
then since we know $Q \vdash \neg Prvbl_Q(\langle \underline{\sigma_Q} \rangle)$
$\Leftrightarrow \sigma_Q$

we could deduce $Q \vdash \sigma_Q$
contradicting $Q \nvdash \sigma_Q$
- Hence $Q \nvdash$ "Q is consistent"
↳ How can we formalize this "proof"?

↳ we'll need to assume more about Q

<u>Def'n</u> Q satisfies the Hilbert–Bernays provability conditions if for all fmlas φ, ψ:

true when Q rec. repr. rec. fns
"can formalize ① in Q"

① if $Q \vdash \varphi$ then $Q \vdash Prvbl_Q(\underline{\langle\varphi\rangle})$
② $Q \vdash [Prvbl_Q(\underline{\langle\varphi\rangle}) \Rightarrow Prvbl_Q(\underline{\langle Prvbl_Q(\underline{\langle\varphi\rangle})\rangle})]$
③ $Q \vdash [Prvbl_Q(\underline{\langle\varphi\rangle}) \wedge Prvbl_Q(\underline{\langle\varphi \Rightarrow \psi\rangle}) \Rightarrow Prvbl_Q(\underline{\langle\psi\rangle})]$

- We write "Q is consistent" as follows
- Let $\bot$ denote $\exists x(x \neq x)$
- Let $Con(Q)$ denote $\neg Prvbl_Q(\underline{\bot})$

<u>Lemma</u> Sps Q satisfies the HB conditions. Then for any φ we have:
$$Q \vdash Con(Q) \Leftrightarrow [\neg Prvbl_Q(\underline{\langle\varphi\rangle}) \vee \neg Prvbl_Q(\underline{\langle\neg\varphi\rangle})]$$

<u>Pf.</u> - $\forall x\,(x = x)$ is a logical axiom
- can deduce $\neg\neg\forall x(x = x)$
- i.e. $\neg\exists x(x \neq x)$
- i.e. $\neg\bot$
- So $Q \vdash \neg\bot$
- Hence by HB(1)
$$Q \vdash Prvbl_Q(\underline{\langle\neg\bot\rangle})$$

(*) <u>Thus</u>: $Q \vdash \neg Con(Q) \Rightarrow [Prvbl_Q(\underline{\langle\bot\rangle}) \wedge Prvbl_Q(\underline{\langle\neg\bot\rangle})]$

Why: left conjunct is immediately deducible from $\neg Con(Q)$, right is above.

- Now: $\bot \Rightarrow (\neg\bot \Rightarrow \varphi)$ is an axiom of our deduction system
- hence by HB ①:

(Δ) $Q \vdash Prvbl_Q(\underline{\langle \bot \Rightarrow (\neg\bot \Rightarrow \varphi)\rangle})$

- Applying HB③ twice we get:

$Q \vdash \neg Con(Q) \Rightarrow Prvbl_Q(\underline{\langle\varphi\rangle})$

Why: $Q \vdash \neg Con(Q) \Rightarrow Prvbl_Q(\underline{\langle\bot\rangle})$ by above

- combining w/ (Δ) and using HB③ get

$Q \vdash \neg Con(Q) \Rightarrow Prvbl_Q(\underline{\langle\neg\bot \Rightarrow \varphi\rangle})$

- but also $Q \vdash \neg Con(Q) \Rightarrow Prvbl_Q(\underline{\langle\neg\bot\rangle})$
- Hence again by HB③ $Q \vdash \neg Con(Q) \Rightarrow Prvbl_Q(\underline{\langle\varphi\rangle})$

By a similar arg get

$Q \vdash \neg Con(Q) \Rightarrow Prvbl_Q(\underline{\langle\neg\varphi\rangle})$

And combining:

$Q \vdash \neg Con(Q) \Rightarrow (Prvbl_Q(\underline{\langle\varphi\rangle}) \wedge Prvbl_Q(\underline{\langle\neg\varphi\rangle}))$

which gives ($\Leftarrow$) direction of claim

- Conversely: by HB ① again $Q \vdash Prvbl_Q(\underline{\langle\varphi \Rightarrow (\neg\varphi \Rightarrow \bot)\rangle})$
- Hence by HB③ twice again:

$Q \vdash Prvbl_Q(\underline{\langle\varphi\rangle}) \wedge Prvbl_Q(\underline{\langle\neg\varphi\rangle}) \Rightarrow Prvbl_Q(\underline{\bot})$ ← i.e. $\neg Con(Q)$

which is ($\Rightarrow$) direction ✓

Theorem (Gödel's 2nd Incompleteness Theorem) Sps Q satisfies HB conditions, represents every recursive function, and is consistent. Then $Q \nvdash Con(Q)$

Pf: Let σ_Q be the Gödel sentence of Q so that

(*) $Q \vdash \neg\sigma_Q \Leftrightarrow Prvbl_Q(\underline{\langle\sigma_Q\rangle})$

- By HB①

(**) $Q \vdash Prvbl_Q(\neg\sigma_Q \Leftrightarrow Prvbl_Q(\underline{\langle\sigma_Q\rangle}))$

- By HB②

$Q \vdash Prvbl_Q(\underline{\langle\sigma_Q\rangle}) \Rightarrow Prvbl_Q(\underline{Prvbl_Q(\underline{\langle\sigma_Q\rangle})})$

- So by HB③ and ($\Leftarrow$) direction of (**):

$Q \vdash Prvbl_Q(\underline{\langle\sigma_Q\rangle}) \Rightarrow Prvbl_Q(\underline{\langle\neg\sigma_Q\rangle})$

- Now using $\Rightarrow$ direction of (*):

$Q \vdash \neg\sigma_Q \Rightarrow (Prvbl_Q(\underline{\langle\sigma_Q\rangle}) \wedge Prvbl_Q(\underline{\langle\neg\sigma_Q\rangle}))$

- by the Lemma $Q \vdash \neg\sigma_Q \Rightarrow \neg Con(Q)$

hence $Q \vdash Con(Q) \Rightarrow \sigma_Q$

Hence if $Q \vdash Con(Q)$ would have $Q \vdash \sigma_Q$

but we know $Q \nvdash \sigma_Q$

Hence $Q \nvdash Con(Q)$ ✓

- So next question is: which Q's satisfy HB conditions?
- Turns out Q_0 not strong enough

Aside. Q_0 is surprisingly weak in other ways, e.g. $Q_0 \vdash \underline{m}+\underline{n} = \underline{n}+\underline{m}$ $\forall n, m \in \mathbb{N}$, but doesn't prove $\forall x \forall y (x+y = y+x)$

↳ it follows: there are nonstandard models $\mathcal{A} \models Q_0$ in which $+$ is not commutative for some nonstd. el'ts of $\mathcal{A}$.

- Despite this, $Q_0 \nvdash Con(Q_0)$ (the above pf doesn't work) but there are closely related systems that do prove their own consistency.

Def'n Peano Arithmetic (PA) is the system obtained by adding induction axioms to PA

i.e. for every fmla $\varphi(x_1, .. x_n, z)$ we add the fmla:

$$\forall x_1 .. \forall x_n [\varphi(\vec{x}, \bar{0}) \wedge [\forall z (\varphi(\vec{x}, z) \Rightarrow \varphi(\vec{x}, z+1))] \Rightarrow \forall z \varphi(\vec{x}, z)]$$

Note: PA is recursive, since one can computably recognize if a given fmla is an induction axiom.

- PA is an infinite axiom scheme, Q_0 is finite
- Can be shown: PA is not finitely axiomatizable.

<u>Theorem</u> PA satisfies the HB provability conditions. Hence (assuming PA is consistent) PA $\nvdash$ Con(PA)

The same is true for any consistent recursive extension of PA

<u>Proof</u> omitted (it's long!)

<u>Remark</u>: It follows that PA $\cup \{\neg \text{Con}(PA)\}$ is consistent! By the completeness theorem, it has a model $\mathcal{A}$.

What does this mean?

$\neg$ Con(PA) asserts "there is an x coding a proof of $\perp$"

$\hookrightarrow$ this does <u>not</u> imply that for some $n \in \mathbb{N}$, "$\underline{n}$ codes a proof of $\perp$" holds in $\mathcal{A}$ (as then n would code an actual proof of $\perp$ from PA)

$\hookrightarrow$ rather, in $\mathcal{A}$, there is a nonstd number N ($> \underline{n}^{\mathcal{A}}$ for all $n \in \mathbb{N}$) that "codes" a "proof" of $\perp$.