

Le second théorème d'incomplétude

Nous clôturons ce livre avec le second théorème d'incomplétude de Gödel.

1. Le second théorème de Gödel

1.1. Énoncé

Théorème 1.

Soit $\mathcal{T}$ une théorie cohérente, contenant $\mathcal{PA}$, ayant des axiomes rékursifs. Alors $\mathcal{T}$ ne peut pas prouver sa propre cohérence.

Si l'on note $\mathcal{E}\text{st-cohérente}(\mathcal{T})$ le prédicat « être une théorie cohérente », la conclusion du théorème est donc :

$$\mathcal{T} \not\vdash \mathcal{E}\text{st-cohérente}(\mathcal{T})$$

Noter que les hypothèses sont les mêmes que celles du premier théorème.

1.2. Cohérence

On rappelle qu'une théorie est *incohérente* si elle prouve n'importe quelle formule, en particulier la formule absurde « $0 = 1$ », ce qui peut s'exprimer dans le langage $\mathcal{L}_0$ par $\mathcal{E}\text{st-démontrable}(\ulcorner 0 = 1 \urcorner)$, où on note bien sûr $1 = S(0)$. Ainsi on définit la formule :

$$\mathcal{E}\text{st-cohérente}(\mathcal{T}) : \neg \mathcal{E}\text{st-démontrable}(\ulcorner 0 = 1 \urcorner)$$

On peut ainsi exprimer la cohérence par une formule de $\mathcal{L}_0$.

2. Preuve

La preuve est une conséquence assez directe du premier théorème si on accepte sa version formalisée.

2.1. Formalisation du premier théorème d'incomplétude

On suppose que $\mathcal{T}$ est une théorie qui contient $\mathcal{PA}$, on suppose aussi que les axiomes de $\mathcal{T}$ sont rékursifs, et on s'intéresse seulement à l'hypothèse de cohérence.

On rappelle que le lemme de diagonalisation fournit une formule $\mathcal{G}$ telle que :

$$\mathcal{T} \vdash (\mathcal{G} \leftrightarrow \neg \mathcal{E}\text{st-démontrable}(\ulcorner \mathcal{G} \urcorner))$$

Et l'énoncé du premier théorème d'incomplétude est alors :

$\mathcal{T}$ cohérente implique $\mathcal{G}$ non démontrable.

Mais par le lemme de diagonalisation, $\mathcal{G}$ signifie exactement « $\mathcal{G}$ non démontrable », donc l'énoncé du premier théorème est aussi :

$\mathcal{T}$ cohérente implique $\mathcal{G}$.

Voici la formalisation du premier théorème d'incomplétude :

$$\boxed{\mathcal{T} \vdash (\text{Est-cohérente}(\mathcal{T}) \rightarrow \mathcal{G})}$$

C'est exactement l'énoncé précédent qui a été *internalisé*. L'énoncé « $\mathcal{T}$ cohérente implique $\mathcal{G}$ » est prouvable (ce qu'on a vu dans les chapitres précédents) mais ici on admet qu'on pourrait en écrire une preuve dans la théorie $\mathcal{T}$. On reviendra sur cette formalisation dans la dernière section.

2.2. Preuve du second théorème d'incomplétude

On admet la version formalisée du premier théorème afin de prouver le second. En cela, on suit la même démarche que Gödel quand il énonce le second théorème et esquisse la preuve à la fin de son article fondateur sur ses théorèmes d'incomplétude.

On a aussi besoin d'admettre le principe appelé « nécessité de la preuve » :

$$\text{si } \mathcal{T} \vdash Q \text{ alors } \mathcal{T} \vdash \text{Est-démontrable}(\ulcorner Q \urcorner)$$

Cela signifie que si une théorie prouve une formule, alors elle sait qu'elle peut prouver cette formule !

Prouvons maintenant qu'une théorie cohérente ne peut pas prouver sa propre cohérence.

On part d'une théorie quelconque $\mathcal{T}$ et on suppose qu'elle prouve sa propre cohérence :

$$\mathcal{T} \vdash \text{Est-cohérente}(\mathcal{T})$$

Par *modus ponens* appliqué à la formalisation du premier théorème d'incomplétude, on a alors :

$$\begin{aligned} & \mathcal{T} \vdash \mathcal{G} \\ \text{donc } & \mathcal{T} \vdash \text{Est-démontrable}(\ulcorner \mathcal{G} \urcorner) \text{ par la nécessité de la preuve} \\ \text{donc } & \mathcal{T} \vdash \neg \mathcal{G} \text{ par l'équivalence de la diagonalisation} \end{aligned}$$

Bilan :

$$\mathcal{T} \vdash \mathcal{G} \quad \text{et} \quad \mathcal{T} \vdash \neg \mathcal{G}$$

Donc $\mathcal{T}$ est incohérente.

Ainsi, par contraposition, on a bien démontré le second théorème :

$$\text{Si } \mathcal{T} \text{ est cohérente alors } \mathcal{T} \not\vdash \text{Est-cohérente}(\mathcal{T}).$$

2.3. Retour sur la formalisation

La preuve expliquée ci-dessus est l'esquisse originale de Gödel (1931). Il faudra attendre les travaux de Hilbert et Bernays (1934-1939), puis Löb (1955) pour obtenir une preuve détaillée. Cette preuve repose sur trois conditions, appelées « conditions de Hilbert-Bernays-Löb » :

- C1 - *Nécessité de la preuve*.

$$\text{Si } \mathcal{T} \vdash Q, \text{ alors } \mathcal{T} \vdash \text{Est-démontrable}(\ulcorner Q \urcorner)$$

- C2 - *Modus ponens des preuves*.

$$\mathcal{T} \vdash \text{Est-démontrable}(\ulcorner \mathcal{P} \rightarrow Q \urcorner) \rightarrow (\text{Est-démontrable}(\ulcorner \mathcal{P} \urcorner) \rightarrow \text{Est-démontrable}(\ulcorner Q \urcorner))$$

- **C3** - *Introspection de la preuve.*

$$\mathcal{T} \vdash \text{Est-démontrable}(\ulcorner \mathcal{Q} \urcorner) \rightarrow \text{Est-démontrable}(\ulcorner \text{Est-démontrable}(\ulcorner \mathcal{Q} \urcorner) \urcorner)$$

On admet que ces conditions sont vérifiées. (Elles sont très techniques à démontrer, il faudra attendre 1955 pour en achever la formalisation.) On peut maintenant donner une preuve rigoureuse de la formalisation du premier théorème d'incomplétude. On part de l'équivalence que la diagonalisation fournit :

$$\mathcal{T} \vdash \mathcal{G} \leftrightarrow \neg \text{Est-démontrable}(\ulcorner \mathcal{G} \urcorner)$$

donc $\mathcal{T} \vdash \mathcal{G} \rightarrow \neg \text{Est-démontrable}(\ulcorner \mathcal{G} \urcorner)$ on n'a gardé que l'implication directe

donc $\mathcal{T} \vdash \text{Est-démontrable}(\ulcorner \mathcal{G} \rightarrow \neg \text{Est-démontrable}(\ulcorner \mathcal{G} \urcorner) \urcorner)$ par (C1)

donc $\mathcal{T} \vdash \text{Est-démontrable}(\ulcorner \mathcal{G} \urcorner) \rightarrow \text{Est-démontrable}(\ulcorner \neg \text{Est-démontrable}(\ulcorner \mathcal{G} \urcorner) \urcorner)$ par (C2)

Mais par (C3), on a toujours :

$$\mathcal{T} \vdash \text{Est-démontrable}(\ulcorner \mathcal{G} \urcorner) \rightarrow \text{Est-démontrable}(\ulcorner \text{Est-démontrable}(\ulcorner \mathcal{G} \urcorner) \urcorner)$$

En posant :

$$\mathcal{Q} : \text{Est-démontrable}(\ulcorner \mathcal{G} \urcorner)$$

on obtient une double conclusion :

$$\mathcal{T} \vdash \text{Est-démontrable}(\ulcorner \mathcal{G} \urcorner) \rightarrow \text{Est-démontrable}(\ulcorner \mathcal{Q} \urcorner) \wedge \text{Est-démontrable}(\ulcorner \neg \mathcal{Q} \urcorner)$$

donc $\mathcal{T} \vdash \text{Est-démontrable}(\ulcorner \mathcal{G} \urcorner) \rightarrow \text{Est-démontrable}(\ulcorner \mathcal{Q} \wedge \neg \mathcal{Q} \urcorner)$

donc $\mathcal{T} \vdash \text{Est-démontrable}(\ulcorner \mathcal{G} \urcorner) \rightarrow \text{Est-démontrable}(\ulcorner 0 = 1 \urcorner)$

ainsi $\mathcal{T} \vdash \text{Est-démontrable}(\ulcorner \mathcal{G} \urcorner) \rightarrow \neg \text{Est-cohérente}(\mathcal{T})$

Par contraposition :

$$\mathcal{T} \vdash \text{Est-cohérente}(\mathcal{T}) \rightarrow \neg \text{Est-démontrable}(\ulcorner \mathcal{G} \urcorner)$$

Mais par diagonalisation, la conclusion est exactement $\mathcal{G}$, donc :

$$\mathcal{T} \vdash \text{Est-cohérente}(\mathcal{T}) \rightarrow \mathcal{G}$$

Ce qui est la formalisation souhaitée du premier théorème d'incomplétude.