From Hilbert to Gentzen and beyond

Reinhard Kahle

Theorie und Geschichte der Wissenschaften
Universität Tübingen and
CMA, FCT, Universidade Nova de Lisboa

Hilbert Bernays Summer School 2019
Tübingen

Partially funded by the FCT projects *Hilbert's 24th Problem*, PTDC/MHC-FIL/2583/2014, and UID/MAT/00297/2019, and the Udo-Keller-Stiftung.

Cantor's Naive Set Theory

Cantor 1895

"By a *set* we understand every collection to a whole M of definite, well-differentiated objects m of our intuition or our thought."

$$M = \{x|\varphi(x)\}, \qquad m \in M \Leftrightarrow \varphi(m)$$

Russell's Paradox

Russell 1901

$$R = \{x | x \notin x\}$$

$$R \overset{?}{\in} R$$

$$R \in R \Leftrightarrow R \in \{x | x \notin x\}$$
$$\Leftrightarrow R \notin R$$
$$\lightning$$

- Historical Note: Zermelo found independently the same paradox: B. Rang and W. Thomas, *Zermelo's discovery of the 'Russell Paradox'*, Historia Mathematica 8(1), 1981, pp. 15–22.

Hilbert's Concerns

$$M = \{x | \varphi(x)\}, \qquad m \in M \Leftrightarrow \varphi(m)$$

- Which $\varphi(x)$ are allowed for meaningful (consistent) set formations?
- Cantor considered the paradoxes as *reductio-ad-absurdum* arguments for the non-existence of a set associated to the underlying "set formations".
- Hilbert—as Frege—was not happy with this "a posteriori view".

Hilbert ~1905

Why is the totality of all sets not permissible?
Why is the set of all real numbers a permissible collection?

- Zermelo's axiomatization appears to be one answer to Hilbert's questions—but it doesn't really answer "Why"!
- Other answers, notably by Poincaré, Weyl, and Brouwer, restrict set theory so far, that certain "usual" mathematical arguments cannot be executed any longer, notably in Analysis.

Hilbert's Programme

- The paradoxes were one of the motivations for Hilbert's Foundational Studies (there are others which, however, we do not address here).
- To secure mathematical reasoning, Hilbert proposed the following strategy for *consistency proofs*:
  1. Formalize mathematical reasoning (proofs).
  2. Showing that no formalized proof can end in a false formula (as, for instance, $0 = 1$).
- Apparently, this is a purely combinatorial question: proofs can be represented by certain sequences of formulas, constructed by clear defined rules, and all one would have to show is, that such a sequence could never have a particular formula as last element.

Note

Hilbert is, by no means, a *formalist* who considers Mathematics as a game with formulas. Formal proofs are just *representation* of "normal mathematical proofs".

Hilbert's Programme

- Initial "philosophical problem" (Poincaré):
  the methods (in particular, induction) used in a "meta proof" (expressing that $0 = 1$ never could be proven) are those which are at stake—thus, one runs in a vicious circle.

- Solution (suggested to Hilbert by Brouwer in 1909):
  using a "weak" theory—whose consistency is beyond doubt—to prove the consistency of strong theories.

First-order languages

Definition

A *first-order language* $\mathcal{L}$ is a set of symbols which can be divided in the following six (disjunctive) subsets:

- logical symbols: $\{\neg, \wedge, \vee, \rightarrow, \forall, \exists, =\}$;
- constant symbols: $\mathcal{C} \subseteq \{c_i | i \in \mathbb{N}\}$,
- function symbols: $\mathcal{F} \subseteq \{f_i^j | i \in \mathbb{N}, j \in \mathbb{N}, j > 0\}$, where f_i^j is the i-th function symbol of arity j;
- relation symbols $\mathcal{R} \subseteq \{R_i^j | i \in \mathbb{N}, j \in \mathbb{N}\}$, where R_i^j is the i-th relation symbol of arity j;
- variables: $\{x, y, z, w, \ldots, x_0, x_1, x_2, \ldots\}$;
- auxiliary signs: $\{$"(", ")", ",", "."$\}$.

First-order languages

According to the definition, for a concrete first-order language we have only to specify only the sets $\mathcal{C}$, $\mathcal{F}$, and $\mathcal{R}$.

Examples

1. For the language $\mathcal{L}_{\text{PA}}$ of the *Peano arithmetic* we have: $\mathcal{C} = \{0\}$, $\mathcal{F} = \{s, +, \cdot\}$, and $\mathcal{R} = \emptyset$, where s is a unary function symbol for the successor function.
2. The language of *set theory* (without urelements) can be given by $\mathcal{C} = \mathcal{F} = \emptyset$ and $\mathcal{R} = \{\in\}$.

Terms

Definition

The *terms* of $\mathcal{L}$ are defined *inductively* as following:

1. Each variable is a term.
2. Each constant symbol is a term.
3. If $t_1, t_2, \ldots, t_n$ are terms and f^n is a n-ary function symbol $(n > 0)$, then the expression $f^n(t_1, t_2, \ldots, t_n)$ is also a term.

Formulae

Definition

The *formulae* of $\mathcal{L}$ are defined inductively as follows:

1. If t_1 and t_2 are terms, then the expression $t_1 = t_2$ is a formula.
2. If $t_1, t_2, \ldots, t_n$ are terms and R^n is a n-ary relation symbol $(n \geq 0)$, then the expression $R^n(t_1, t_2, \ldots, t_n)$ is a formula.
3. If φ and ψ are formulae, then the following expressions are also formulae:
$$(\neg\varphi),\ (\varphi \wedge \psi),\ (\varphi \vee \psi),\ (\varphi \rightarrow \psi).$$
4. If φ is a formula and x a variable, then the expressions $(\forall x.\varphi)$ and $(\exists x.\varphi)$ are also formulae.

Hilbert-style calculus I

Definition

We define the *Hilbert-style calculus* **H** as a derivation system with the following (logical) axioms and rules:

1. The following formulae are axioms:
   - $\vdash \varphi \to (\psi \to \varphi)$
   - $\vdash (\varphi \to (\chi \to \psi)) \to (\varphi \to \chi) \to (\varphi \to \psi)$
   - $\vdash (\neg\varphi \to \neg\psi) \to \psi \to \varphi$
   - $\vdash \varphi \to (\varphi \vee \psi)$
   - $\vdash \psi \to (\varphi \vee \psi)$
   - $\vdash (\varphi \to \chi) \to ((\psi \to \chi) \to (\varphi \vee \psi \to \chi))$
   - $\vdash (\varphi \wedge \psi) \to \varphi$
   - $\vdash (\varphi \wedge \psi) \to \psi$
   - $\vdash \varphi \to (\psi \to (\varphi \wedge \psi))$

Hilbert-style calculus II

Definition

2. Equality axioms.
   - $(u = u)$,
   - $(u = w) \to (w = u)$,
   - $(u_1 = u_2 \wedge u_2 = u_3) \to (u_1 = u_3)$,
   - $(u_1 = w_1 \wedge \cdots \wedge u_n = w_n) \to (R(u_1, \ldots, u_n) \to R(w_1, \ldots, w_n))$,
   - $(u_1 = w_1 \wedge \cdots \wedge u_m = w_m) \to (t[u_1, \ldots, u_m] = t[w_1, \ldots, w_m])$,

   where $u, w, u_1, \ldots$ are variables and constant symbols, R a n-ary relation symbol, and t a term, in which $u_1, \ldots, u_m$ or $w_1, \ldots, w_m$ may occur.
3. Quantifier axioms:
   - $\vdash (\forall x.\varphi(x)) \to \varphi(t)$
   - $\vdash \varphi(t) \to (\exists x.\varphi(x))$

Hilbert-style calculus III

Definition

As rules we have:

4. Modus Ponens.

$$\frac{\vdash \varphi \to \psi \qquad \vdash \varphi}{\vdash \psi}$$

5. Generalisation; let x be a variable not free in φ.

$$\frac{\vdash \varphi \to \psi(x)}{\vdash \varphi \to \forall y.\psi(y)}$$

$$\frac{\vdash \psi(x) \to \varphi}{\vdash (\exists y.\psi(y)) \to \varphi}$$

Proof in **H**

Definition

A *proof of* φ *starting from a set of formulae* Φ (in the Hilbert-style calculus **H**), is a *finite* sequence of formulae $\psi_1, \psi_2, \ldots, \psi_n$ with $\psi_n = \varphi$, and each of these formulae ψ_i is either

- an axiom of **H**,
- an element of Φ, or
- is obtained from the previous formulae ψ_j, $j < i$, by an application of a rule.

We say that φ *is provable from* Φ (in the Hilbert-style calculus **H**), and write $\Phi \vdash \varphi$, if there exists a proof of φ starting from Φ.

Example

$\varphi \to \varphi$ is not an axiom in our calculus.

Beispiel

$\vdash (\varphi \to ((\varphi \to \varphi) \to \varphi)) \to (\varphi \to (\varphi \to \varphi)) \to (\varphi \to \varphi)$	Second axiom
$\vdash \varphi \to ((\varphi \to \varphi) \to \varphi)$	First axiom
$\vdash (\varphi \to (\varphi \to \varphi)) \to (\varphi \to \varphi)$	Modus Ponens
$\vdash \varphi \to (\varphi \to \varphi)$	First axiom
$\vdash \varphi \to \varphi$	Modus Ponens

Peano arithmetic

We use the language of Peano arithmetic $\mathcal{L}_{\text{PA}} = \{0, s, +, \cdot\}$.

Definition (Peano arithmetic)

Peano arithmetic PA comprises the following six non-logical axioms and the following axiom scheme:

- PA1 $\forall x.\neg(s(x) = 0)$,
- PA2 $\forall x, y.s(x) = s(y) \to x = y$,
- PA3 $\forall x.x + 0 = x$,
- PA4 $\forall x, y.x + s(y) = s(x + y)$,
- PA5 $\forall x.x \cdot 0 = 0$,
- PA6 $\forall x, y.x \cdot s(y) = (x \cdot y) + x$.

The axiom scheme of complete induction:

$$\varphi(0) \wedge (\forall y.\varphi(y) \to \varphi(s(y))) \to \forall x.\varphi(x).$$

$\text{PA} \vdash \varphi$ iff there is a finite set Φ of axioms of PA such that $\Phi \vdash \varphi$.

Peano arithmetic

- Hilbert's Programme for PA: showing that $\mathsf{PA} \nvdash 0 = 1$.
- Apparently unrelated question:

Is PA *syntactically complete*, i.e., does for every formula φ holds that:

$$\mathsf{PA} \vdash \varphi \quad \text{or} \quad \mathsf{PA} \vdash \neg\varphi\ ?$$

- Gödel's First Incompleteness theorem shows that this is not the case.
- Gödel's Second Incompleteness theorem shows that the First Incompleteness theorem entails the impossibility of a consistency proof for PA (and all stronger systems) in the way Hilbert had envisaged them.

Gödel's First Incompleteness Theorem

- The first incompleteness theorem shows that the Peano Arithmetic is *syntactically* incomplete. That means, there is a formula φ such that

$$\mathsf{PA} \nvdash \varphi \quad \text{and} \quad \mathsf{PA} \nvdash \neg\varphi.$$

- The idea of the proof is quite simple. Consider the classical paradox of the *liar*:

  This sentence is false.

  Obviously, the sentence can neither be *true* nor *false* without provocating a contradiction.
- In analogy, consider now the following *Gödel sentence*:

  This sentence is not provable.

  If this sentence can be represented *faithfully* in the language of Peano-Arithmetic, it can neither be provable nor refutable (i.e., its negation would be provable).

Two challenges

To formalize the Gödel sentence “This sentence is not provable.” in PA we have to solve two problems:

1. Formalizing *provability*.
2. Expressing the self-reference (“*This* sencence ...”).

The proof predicate

- Formulas are strings of symbols, which can be coded by numbers, its *Gödel number*:

$$\varphi \mapsto \ulcorner\varphi\urcorner \in \mathbb{N}.$$

- Proofs are finite sequences of formulas (obeying the derivation rules of the calculus); thus, a proof can be coded by a sequence of the corresponding Gödel numbers:

$$\langle \ulcorner\varphi_1\urcorner, \ulcorner\varphi_2\urcorner, \dots, \ulcorner\varphi_n\urcorner \rangle \in \mathbb{N}.$$

- All this coding can be done within the realm of primitive recursive functions.
- With some technical work, one can define a primitive recursive relation $\mathrm{Bew}_{\mathsf{PA}}$ such that $\mathrm{Bew}_{\mathsf{PA}}(x, y)$ is true, if and only if x is the Gödel number of a proof in PA of the formula with the Gödel number y.

Representability

Let $\bar{n}$ is a term of the language of the formal theory T representing the natural number n.

Definition

Let T be an arbitrary theory.

- A relation $R \subseteq \mathbb{N}^n$ is *numeralwise representable* in T by a formula φ if one has, for all natural numbers $m_1, \ldots, m_n$:
  $R(m_1, \ldots, m_n)$ is true if and only if $\mathsf{T} \vdash \varphi(\bar{m}_1, \ldots, \bar{m}_n)$,
  We also say φ *numerates* the relation R in T.
- φ *binumerates* R in T if it numerates it and one has also:
  $R(m_1, \ldots, m_n)$ is false if and only if $\mathsf{T} \vdash \neg\varphi(\bar{m}_1, \ldots, \bar{m}_n)$.

Representability

Theorem (Representation Theorem)

PA binumerates all primitive-recursive relations.

This theorem applies to $\mathrm{Bew}_{\mathsf{PA}}$ and we have that there is a formula $\mathrm{Bew}_{\mathsf{PA}}$ in the language of PA with:

$$\mathrm{Bew}_{\mathsf{PA}}(m_1, m_2) \text{ is true if and only if } \mathsf{PA} \vdash \mathrm{Bew}_{\mathsf{PA}}(\bar{m}_1, \bar{m}_2)$$

$$\mathrm{Bew}_{\mathsf{PA}}(m_1, m_2) \text{ is false if and only if } \mathsf{PA} \vdash \neg\mathrm{Bew}_{\mathsf{PA}}(\bar{m}_1, \bar{m}_2).$$

A provability predicate

- By definition of the relation $\mathrm{Bew}_{\mathsf{PA}}$ we have for its representation $\mathrm{Bew}_{\mathsf{PA}}$ in PA:

$$\begin{aligned} \mathsf{PA} \vdash \varphi &\iff \mathsf{PA} \vdash \mathrm{Bew}_{\mathsf{PA}}(t, \ulcorner\varphi\urcorner) \qquad \text{for a closed term } t \\ &\implies \mathsf{PA} \vdash \exists x.\mathrm{Bew}_{\mathsf{PA}}(x, \ulcorner\varphi\urcorner) \\ &\iff \mathsf{PA} \vdash \mathrm{B}_{\mathsf{PA}}(\ulcorner\varphi\urcorner) \end{aligned}$$

  t is a sequence number of $\langle \ulcorner\varphi_0\urcorner, \ulcorner\varphi_1\urcorner, \ldots, \ulcorner\varphi_{n-1}\urcorner, \ulcorner\varphi\urcorner \rangle$.

- In short:

$$\mathsf{PA} \vdash \varphi \implies \mathsf{PA} \vdash \mathrm{B}_{\mathsf{PA}}(\ulcorner\varphi\urcorner) \tag{1}$$

- Note that we don't have immediately the "missing" direction:

$$\mathsf{PA} \vdash \exists x.\mathrm{Bew}_{\mathsf{PA}}(x, \ulcorner\varphi\urcorner) \implies \mathsf{PA} \vdash \mathrm{Bew}_{\mathsf{PA}}(t, \ulcorner\varphi\urcorner)$$

- In general, one cannot conclude from an existential statement like $\exists x.\mathrm{Bew}_{\mathsf{PA}}(x, \ulcorner\varphi\urcorner)$ that there is also a *closed term* which exemplifies such an x.

Diagonalization lemma

Theorem (Diagonalization lemma)

Let $\varphi(x)$ be a formula with exactly one free variable x. Then there is a sentence ψ such that:

$$\mathsf{PA} \vdash \psi \leftrightarrow \varphi(\ulcorner\psi\urcorner).$$

Proof.

Define $\vartheta(x)$ as $\varphi(\mathrm{Sub}(x, \mathrm{Num}(x)))$. Let $\bar{m}$ be $\ulcorner\vartheta(x)\urcorner$ and let ψ be $\vartheta(\bar{m})$.

$$\begin{aligned} \psi &\leftrightarrow \vartheta(\bar{m}) \\ &\leftrightarrow \varphi(\mathrm{Sub}(\bar{m}, \mathrm{Num}(\bar{m}))) \\ &\leftrightarrow \varphi(\mathrm{Sub}(\ulcorner\vartheta(x)\urcorner, \ulcorner\bar{m}\urcorner)) \\ &\leftrightarrow \varphi(\ulcorner\vartheta(\bar{m})\urcorner) \\ &\leftrightarrow \varphi(\ulcorner\psi\urcorner) \end{aligned}$$

ψ expresses "I have the property φ".