Werk

Label: Article

Jahr: 1972

PURL: https://resolver.sub.uni-goettingen.de/purl?316342866_0013|log45

Kontakt/Contact

ON INTERPRETABILITY IN SET THEORIES   II

Petr HÁJEK, Praha

This paper is a continuation of [2] and [3] and uses techniques developed in [1]. ZF denotes the Zermelo-Fraenkel set theory and GB the Gödel-Bernays set theory. We adopt conventions made in [3] § 1 (Preliminaries). GB is a conservative extension of ZF; so we have $Con\,(ZF, \varphi) \Longleftrightarrow$ $\Longleftrightarrow Con\,(GB, \varphi)$ for each ZF-formula φ. Denote by $\mathcal{I}_{ZF}$ ($\mathcal{I}_{GB}$) the set of all ZF-formulas φ such that (ZF, φ) is interpretable in ZF $((GB, \varphi)$ is interpretable in GB). We know the following: (1) $\varphi \in \mathcal{I}_{ZF} \cup$ $\cup \mathcal{I}_{GB} \Rightarrow Con\,(ZF, \varphi)$, (2) $\mathcal{I}_{ZF} - \mathcal{I}_{GB} \neq \emptyset$, (3) $\mathcal{I}_{ZF} \in \Pi_2^0 -$ $- \Sigma_1^0$ and $\mathcal{I}_{GB} \in \Sigma_1^0$.

(We assume $Con\,(ZF)$.) There remain the following questions:

(1) What is the exact position of $\mathcal{I}_{ZF}$ in the arithmetical hierarchy? In particular, is $\mathcal{I}_{ZF}$ a complete Π_2^0-set?

(2) What is the relation between $Con\,(ZF, \varphi)$, $\varphi \in$ $\in \mathcal{I}_{ZF}$, $\varphi \in \mathcal{I}_{GB}$? In particular, is $\mathcal{I}_{GB} - \mathcal{I}_{ZF}$ nonempty?

Unfortunately, I have not succeeded to answer these questions exhaustively; but I hope that the results of this paper give some new information on both questions. We prove the following:

Theorem 1. If ZF is consistent then $\mathcal{J}_{ZF} \notin \Pi_1^0$.

The question if $\mathcal{J}_{ZF}$ is not a Σ_2^0 -set, in particular, if it is a complete Π_2^0 -set, remains open. According to question (2), if we had a (closed) formula $\varphi \in \mathcal{J}_{GB} - \mathcal{J}_{ZF}$, it would satisfy the following: $Con\,(ZF,\varphi)$, $Con\,(ZF, \neg\varphi)$ (i.e. φ would be independent from ZF), $\varphi \notin \mathcal{J}_{ZF}$. I offer to the reader a formula with the following properties:

Theorem 2. If ZF is consistent then there is a closed ZF-formula φ such that (1) φ is independent from ZF, (2) neither (ZF, φ) nor $(ZF, \neg\varphi)$ is interpretable in ZF and (3) neither (GB, φ) nor $(GB, \neg\varphi)$ is interpretable in GB.

In Discussion, we mention possible generalizations of these results (in the spirit of [3]) for theories containing arithmetic and having some additional properties; we further show that if $\mathcal{J}_{GB} - \mathcal{J}_{ZF}$ is non-empty then there is a very simple formula in this set. We conclude with some remarks.

It seems reasonable to use the following hierarchy of P-formulas (P is the Peano arithmetic): a P-formula is Π_m (Σ_m) if it has a prefix containing m alterating quantifiers, the first one being universal (existential), followed by a PR-formula (see [1] for PR-formulas). There will

be no misunderstanding with the arithmetical hierarchy of sets of natural numbers (here we use Σ_m^o and Π_m^o). If Γ is a class of formulas and T is a theory then we say that a T-formula φ is a Γ -formula in T if there is a Γ -formula ψ such that $T \vdash \varphi \equiv \psi$. Note that for each T containing P Σ_1' -formulas in T coincide with (Feferman's) RE-formulas in T .

Lemma 1. If $*$ is an interpretation of ZF in ZF then there is a formula φ with two free variables such that the following is ZF-provable ($x, y, \ldots$ are variables for natural numbers and $x^*, y^*, \ldots$ are variables for natural numbers in the sense of the interpretation):

(1) $(\forall x)(\exists ! x^*)\varphi(x, x^*)$,

(2) $\varphi(\overline{0}, \overline{0}^*)$

(3) $(\varphi(x, x^*) \& \varphi(x + \overline{1}, y^*)) \to y^* = x^* +^* \overline{1}^*$.

Proof. Let $Seq(a)$ mean that a is a finite sequence, let $\ell h(a)$ be the length of the sequence and let $(a)_i$ be the i-th member of a . We put

$$\varphi(x, x^*) \equiv (\exists a)(Seq(a) \& \ell h(a) = x + \overline{1} \& (a)_{\overline{0}} =$$
$$= \overline{0}^* \& (\forall y < \ell h(a) \dot- \overline{1})((a)_{i+\overline{1}} = (a)_i +^* \overline{1}^*)) .$$

One proves the above formulas by induction inside ZF.

Lemma 2. If $*$ is an interpretation of ZF in ZF and if φ is as in Lemma 1 then for each Σ_1 -formula $\varphi(x, \ldots)$ we have:

(*) $ZF \vdash (\varphi(x, x^*) \& \ldots) \to (\varphi(x, \ldots) \to \varphi^*(x^*, \ldots))$.

<u>Proof</u>. By [1] 3.9, it suffices to prove the present lemma for Feferman's BPF. First one proves by (metamathematical) induction

$$ZF \vdash (\varphi(x, x^*) \& \ldots) \to (\psi(x, \ldots) \equiv \psi^*(x^*, \ldots))$$

for each $\psi \in EF$ using induction inside ZF; then one proves ($*$) for BPF (derive the following formulas from (1) - (3) in ZF :

(4) $\quad (\varphi(x, u^*) \& \varphi(y, u^*)) \to x = y$,

(5) $\quad (\varphi(x, u^*) \& v^* <^* u^*) \to (\exists y < x) \varphi(y, v^*))$.

<u>Corollary 1</u>. If φ is a PR-formula then
$$(\varphi(x, x^*) \& \ldots) \to (\varphi(x, \ldots) \equiv \varphi^*(x^*, \ldots))$$
(since both φ and $\neg \varphi$ are Σ_1'-formulas in ZF).

<u>Corollary 2</u>. If φ is a Π_1-formula then
$$(\varphi(x, x^*) \& \ldots) \to (\varphi^*(x^*, \ldots) \to \varphi(x, \ldots))$$

<u>Corollary 3</u>. If φ is a closed Π_1-formula and $\varphi \in \mathfrak{I}_{ZF}$ then $ZF \vdash \varphi$.

It is of some interest that we can give an alternative proof of the last corollary using the Orey's result (cf. [3] Lemma 2):

Let k be such that all the axioms of the arithmetic Q are provable in $ZF \vdash k$. Then

$$ZF \vdash \neg \varphi \to \mathcal{P}r_{[Q]}(\neg \overline{\varphi}) \to \mathcal{P}r_{[ZF \vdash k]}(\neg \overline{\varphi}) \to \neg Con_{[(ZF \vdash k, \varphi)]} ,$$

i.e. $ZF \vdash Con_{[(ZF \vdash k, \varphi)]} \to \varphi$, which together with

Orey's result gives the corollary. (For the first impli-
cation see [1] 5.5.)

Lemma 3 (Feferman [1] 6.6 and 8.9). If ξ is a PR-
bi-numeration of ZF then $(\neg \, Con_\xi) \in \mathcal{J}_{ZF}$.

Proof of Theorem 1. Suppose that $\mathcal{J}_{ZF}$ is Π_1^0 ,
i.e. the complement of $\mathcal{J}_{ZF}$ is recursively enumerable.
Let ξ be a PR-bi-numeration of ZF in ZF; then
$(ZF, \neg \, Con_\xi)$ is consistent and, by [3] Lemma 1, there
is a "nice" numeration of $- \mathcal{J}_{ZF}$ in $(ZF, \neg \, Con_\xi)$, i.e.
there is a P-formula γ such that

$$\varphi \notin \mathcal{J}_{ZF} \Longleftrightarrow (ZF, \neg \, Con_\xi) \vdash (\exists y) \, \gamma \, (\overline{\varphi}, y) \Longleftrightarrow$$

$$\Longleftrightarrow (\exists k)((ZF, \neg \, Con_\xi) \vdash \gamma \, (\overline{\varphi}, \overline{k})) \; .$$

Note that $ZF \vdash \gamma \, (\overline{\varphi}, y) \equiv \alpha_0 \, (\overline{f(\varphi)}, y)$ where α_0 is a
Π_1 -formula in ZF defined in [3] and f is a recursive
function; hence $\gamma \, (\overline{\varphi}, y)$ is a Π_1 -formula in ZF.

If φ is a formula and $\varphi \notin \mathcal{J}_{ZF}$ then (i)
$ZF \nvdash \varphi$, (ii) for some k we have
$ZF, \neg \, Con_\xi \vdash \gamma \, (\overline{\varphi}, \overline{k})$ and by (i) we have
(iii) $ZF \vdash (\forall y < \overline{k}) \neg \, Pr \, f_\xi \, (\overline{\varphi}, y)$.

By the diagonal lemma [1] 5.1, find a closed P-formula
such that

$$ZF \vdash \varphi \equiv (\exists y) \, (\gamma \, (\overline{\varphi}, y) \, \& \, (\forall z < y) \neg \, Pr \, f \, (\overline{\varphi}, z)) \; .$$

Suppose $\varphi \notin \mathcal{J}_{ZF}$; then, by (ii) and (iii) above,
we have $ZF, \neg \, Con_\xi \vdash \varphi$. Since $\neg \, Con_\xi \in \mathcal{J}_{ZF}$ (by
Lemma 3), we have $\varphi \in \mathcal{J}_{ZF}$.

Suppose $\varphi \in \mathcal{I}_{ZF}$, then

$Con(ZF, \neg Con_{\xi}, (\forall y) \neg \gamma(\overline{\varphi}, y))$ by the properties of

γ . Denote the interpretation of (ZF, φ) in ZF by

$*$ and the theory $(ZF, \neg Con_{\xi}, (\forall y) \neg \gamma(\overline{\varphi}, y)$ by ZF_1 .

Then we have

(1) $ZF_1 \vdash (\exists x) \, Prf_{\xi}(\overline{\varphi}, x)$ (from $\neg Con_{\xi}$)

(2) $ZF_1 \vdash \neg(\exists y) \gamma(\overline{\varphi}, y)$

(3) $ZF_1 \vdash (\exists y^*)(\gamma^*(\overline{\varphi}^*, y^*) \& (\forall x^* <^* y^*) \neg Prf^*_{\xi*}(\overline{\varphi}^*, x^*))$.

We proceed informally in ZF_1 . Let φ be as in Lem-

ma 1. For y^* from (3), there is no y such that

$\varphi(y, y^*)$ (say, y^* is non-standard); otherwise we had

$\gamma(\overline{\varphi}, y)$ by Corollary 2. But if x is as in (1) and

if $\varphi(x, x^*)$ then $Prf^*_{\xi*}(\overline{\varphi}^*, x^*)$ and neces-

sarily $x^* <^* y^*$ (cf. (5) in the proof of Lemma 2!).

This contradicts (3). So we derived a contradiction in ZF .

Hence we proved $\varphi \notin \mathcal{I}_{ZF}$.

We see that the assumption $\mathcal{I}_{ZF} \in \Pi^0_1$ leads to

a contradiction; hence $\mathcal{I}_{ZF}$ is not Π^0_1 .

Lemma 4 (Vopěnka [4]). $(GB, \neg Con_{[GB]})$ is inter-

pretable in GB, i.e. $(\neg Con_{[GB]}) \in \mathcal{I}_{GB}$.

Proof of Theorem 2. Let $Intp(x, y)$ be a PR-for-

mula saying " y is an interpretation of $([GB], x)$ in

[GB]"(cf. [2] or [3]) and find a φ such that

$$ZF \vdash \varphi \equiv (\forall x)(\mathcal{I}n\mathfrak{.}t\,\varrho\,(\overline{\varphi}, x) \to (\exists y < x)\mathcal{I}n\mathfrak{.}t\varrho\,(\neg\overline{\varphi}, y))$$

(by the way, φ is the Rosser's formula with interpretability instead of provability).

(1) Let d be the least interpretation of (GB, φ) in GB; denote it by $*$. Then $GB \vdash \varphi^*$,

$GB \vdash (\mathcal{I}n\mathfrak{.}t\,\varrho\,(\overline{\varphi}, \overline{d}))^*$, i.e.

$GB \vdash [(\exists y < \overline{d})\,\mathcal{I}nt\,\varrho\,(\overline{\neg\varphi}, y)]^*$. The formula

$(\exists y < \overline{d})\,\mathcal{I}n\mathfrak{.}t\,\varrho\,(\overline{\neg\varphi}, y)$ is PR in GB, hence, by Corollary 1, $GB \vdash (\exists y < \overline{d})\,\mathcal{I}n\mathfrak{.}t\,\varrho\,(\overline{\neg\varphi}, \overline{d})$ and hence there is a $d_1 < d$ which is an interpretation of $(GB, \neg\varphi)$ in GB. Denote it by $\square$. We have

$GB \vdash \neg\varphi^\square$, $GB \vdash [\mathcal{I}n\mathfrak{.}t\,\varrho\,(\overline{\neg\varphi}, \overline{d_1})]^\square$ and hence

$GB \vdash [(\exists z < \overline{d_1})\,\mathcal{I}n\mathfrak{.}t\,\varrho\,(\overline{\varphi}, z)]^\square$ and

$GB \vdash (\exists x < \overline{d_1})\,\mathcal{I}n\mathfrak{.}t\,\varrho\,(\overline{\varphi}, x)$, so that there is a $d_2 < d_1 < d$ which is an interpretation of (GB, φ) in GB. This is a contradiction, so that $\varphi \notin \mathcal{I}_{GB}$.

(2) If $(\neg\varphi) \in \mathcal{I}_{GB}$ then there is the least d_1 which is an interpretation of $(GB, \neg\varphi)$ in GB. By (1), then there is a d_2 which is an interpretation of (GB, φ) in GB, which is a contradiction. Hence $(\neg\varphi) \notin \mathcal{I}_{GB}$ and φ is independent from GB (and from ZF).

(3) $\varphi \notin \mathcal{I}_{ZF}$ since φ is a Π_1-formula in ZF (cf. Corollary 3).

- 451 -

(4) To prove $(\neg \varphi) \notin \mathcal{J}_{ZF}$ we need the following

Lemma 5. If ξ is a PR-bi-numeration of ZF such that $ZF \vdash Con_{[GB]} \equiv Con_\xi$ and if φ is as above then $(ZF, \neg Con_\xi) \not\vdash \neg \varphi$.

Otherwise we had the following interpretations:

$$GB, \neg \varphi \Longrightarrow GB, \neg Con_\xi \Longrightarrow GB, \neg Con_{[GB]} \longrightarrow GB \ .$$

(Double arrows are identities; for the last arrow see Lemma 4.) By composition of interpretations, we would have an interpretation of $(GB, \neg \varphi)$ in GB, which is a contradiction. (Note that the "natural" bi-numeration of ZF has the desired property.)

We continue the proof of $(\neg \varphi) \notin \mathcal{J}_{ZF}$. Suppose the contrary. Then we have the following interpretations:

$$ZF, \neg \varphi \longrightarrow ZF \Longrightarrow ZF, \neg Con_\xi, \varphi \ .$$

We consider the composed interpretation $*$ of $(ZF, \neg \varphi)$ in $(ZF, \neg Con_\xi, \varphi)$ and proceed in the last theory. Since $\neg Con_\xi$ we have $\neg Con_{[GB]}$ and hence there are y, x such that $Int_p(\overline{\varphi}, y)$ and $Int_p(\overline{\neg\varphi}, x)$. Suppose that y and x are least with the corresponding properties. Then, by φ, x is smaller than y. On the other hand, we have $(\neg \varphi)^*$, which says

$$(\exists u^*)(Int_p^*(\overline{\varphi}^*, u^*) \ \& \ (\forall v^* <^* u^*) \neg Int_p^*(\overline{\neg\varphi}^*, v^*)) \ .$$

If $\varphi(x, x^*)$ then we have $Int_p^*(\overline{\neg\varphi}^*, x^*)$ and hence $u^* <^* x^*$; then there is a u such that

- 452 -

$\rho(u, u*)$, $u < x$ and $\mathcal{J}_m t \rho(\overline{\varphi}, u)$ which
is a contradiction. Since $(ZF, \neg \text{Con}_\xi, \varphi)$ is consistent
by Lemma 5, there is no interpretation of $(ZF, \neg \varphi)$
in ZF, q.e.d.

Discussion. (1) Let us first discuss the possibility
of generalizing Theorems 1 and 2 for theories containing
arithmetic. Inspection of the proof of Theorem 1 shows that
its assertion holds for any primitive recursively axiomati-
zed theory T containing P, which is consistent, essen-
tially reflexive (so that [3] Lemma 1 applies) and satis-
fies Lemmas 1 and 2. The proofs of these lemmas apply to
each theory T in which, in addition to the assumptions
just made, the induction schema is provable for all T-for-
mulas and in which sequences of arbitrary objects are defi-
nable. (Note in passing that in GB sequences of arbitrary
classes are easily definable, but the induction schema is
not provable for all formulas.) Concerning Theorem 2, let
$P \subseteq T \subseteq S$, where T is as above and S is a conser-
vative finitely axiomatized extension of T. We need two
additional assumptions m S : (i) There is a PR-bi-numera-
tion α of T in T such that $T \vdash \text{Con}_{[S]} \cong \text{Con}_\alpha$.
(This is the case e.g. if the formal statement saying
" $[S]$ is a conservative extension of α " is provab-
le in T .)

(ii) $(S, \neg \text{Con}_{[S]})$ is interpretable in S .
This is an important assumption; it is not clear how to

modify Vopěnka's original proof of $(\neg \; Con_{[GB]}) \notin \mathcal{I}_{GB}$
e.g. for a proof of $\neg \; Con_{[GP]} \notin \mathcal{I}_{GP}$ where GP is the
finitely axiomatized conservative extension of the Peano's
arithmetic with classes (say, Gödel-Peano). Let us stress
the fact that one cannot use Feferman's [1] 8.9 for S sin-
ce S is finitely axiomatizable and therefore not reflexi-
ve.

(2) Suppose that we would find a ZF-formula φ such
that $\varphi \in \mathcal{I}_{GB} - \mathcal{I}_{ZF}$. Then, by Orey's result, there is
a natural number k such that $Con_{[ZF \upharpoonright k, \varphi]}$ is not
provable in ZF. Denote the last formula by φ_0 . It is a
P-formula and, moreover, a π_1 -formula. Since ZF $\nvdash \varphi_0$
we have $\varphi_0 \notin \mathcal{I}_{ZF}$ by Corollary 3. On the other hand,
if $*$ is an interpretation of (GB, φ) in GB then
$GB \vdash \varphi^*$, $GB \vdash (\varphi \rightarrow \varphi_0)^*$ by essential reflexivi
ty of ZF and by $ZF \subsetneq GB$; hence we have $GB \vdash \varphi_0^*$ and
$\varphi_0 \in \mathcal{I}_{GB}$. So we have proved the following

Fact. If $\mathcal{I}_{GB} - \mathcal{I}_{ZF} \neq \emptyset$ then there is a π_1 -for-
mula in $\mathcal{I}_{GB} - \mathcal{I}_{ZF}$.

This contrasts with Corollary 3; by this corollary, no
π_1 -formula is in $\mathcal{I}_{ZF} - \mathcal{I}_{GB}$ (Examples of formulas in
$\mathcal{I}_{ZF} - \mathcal{I}_{GB}$ constructed in [2] and [3] are π_2 -formu-
las.)

(3) It follows by Orey's result that $\varphi \in \mathcal{I}_{ZF}$ iff
there is a recursive function f such that, for each k ,

$f(\ell)$ is a proof of $Con_{[ZF \vdash \ell, \varphi]}$ in ZF. Define $\varphi \in \mathcal{J}_{ZF}^{Prim}$ iff there is a primitive recursive function f such that, for each ℓ, $f(\ell)$ is a proof of $Con_{[ZF \vdash \ell, \varphi]}$ in ZF. Then $\mathcal{J}_{ZF}^{Prim}$ is Σ_2^0 (by the existence of a recursive function universal for primitive recursive functions). Inspection of the proof in [2] shows that $\mathcal{J}_{ZF}^{Prim} - \mathcal{J}_{GB}$ is non-empty (assuming that ZF is ω-consistent).

Is $\mathcal{J}_{ZF} - \mathcal{J}_{ZF}^{Prim} \neq \emptyset$? Can we weaken the assumption of ω-consistency to $Con(ZF)$ in the proof of $\mathcal{J}_{ZF}^{Prim} - \mathcal{J}_{GB} \neq \emptyset$ using methods of [3] or other methods?

R e f e r e n c e s

[1] S. FEFERMAN: Arithmetization of metamathematics in a
        general setting, Fundamenta Mathematicae 49(1966),
        35-92.

[2] P. HÁJEK: On interpretability in set theories, Comment.
        Math.Univ.Carolinae 12(1971),73-79.

[3] M. HÁJKOVÁ, P. HÁJEK: On interpretability in theories
        containing arithmetic, Fundamenta Mathematicae
        LXXVI(1972)(to appear).

[4] P. VOPĚNKA: A new proof of Gödel's result on non-pro-
        vability of consistency, Bull.Acad.Polon.Sci.XIV
        (1966),111-115.

Matematický ústav ČSAV

Žitná 25, Praha 1

Československo