8. GENERALIZATIONS

So far our results have been explicitly stated (and proved) only for theories of first order arithmetic. But, as mentioned in the introduction, they hold, after suitable reformulation, in a much more general setting. Needless to say, we are not going to show this in every detail. In fact, we shall skip Chapters 3, 5, 7 altogether and concentrate on some of the main results of Chapters 2, 4, and 6. These examples should enable the reader to generalize (most of) the results of the preceding chapters.

In this chapter he theories S, T, etc. are no longer arithmetical theories, but they are still consistent and primitive recursive and we assume that the languages of these theories are always finite. L_T is the language of T. T is a *pure* extension of S if $S \dashv T$ and $L_T = L_S$. Lower case Greek letters are now used for formulas of L_T as well as for formulas of L_A.

We assume that the reader can extend the definition of t: $S \leq T$ to the present more general setting. Let $t^{-1}(T) = \{\varphi: T \vdash t(\varphi)\}$. Then $t^{-1}(T) \vdash \psi$ iff $T \vdash t(\psi)$. Since L_S is finite, t is primitive recursive.

The following lemma is immediate.

Lemma 1. (a) t: $S \leq T$ iff $S \dashv t^{-1}(T)$.

(b) t: $t^{-1}(T) \leq T$ and so $t^{-1}(T) \leq T$; in fact, t: $t^{-1}(T) \trianglelefteq T$; it follows that $t^{-1}(T)$ is consistent.

(c) $t^{-1}(T + t(\varphi)) \dashv\vdash t^{-1}(T) + \varphi$.

§1. Incompleteness. Our first result, Gödel's incompleteness theorem, is a straightforward generalization of Theorem 2.1; $\delta_t(x,y)$ is a formula defining t as in Fact 2.

Theorem 1. Suppose t: $Q \leq T$. Let φ be such that

(Gt) $\quad Q \vdash \varphi \leftrightarrow \neg\exists y(\delta_t(\varphi,y) \wedge Pr_T(y))$.

Then φ is a true Π_1 sentence such that $T \nvdash t(\varphi)$. Hence if $t^{-1}(T)$ is Σ_1–sound, then also $T \nvdash \neg t(\varphi)$.

By Theorem 1, for each t: $Q \leq T$, there is a true Π_1 sentence φ_t such that $T \nvdash t(\varphi_t)$. By a similar generalization of Rosser's theorem, we obtain a Π_1 sentence θ_t such that $T \nvdash t(\theta_t)$ and $T \nvdash \neg t(\theta_t)$. This result can be improved by showing that there is a single Π_1 sentence ψ such that $T \nvdash t(\psi)$ and $T \nvdash \neg t(\psi)$ for every t: $Q \leq T$:

Theorem 2. There is a (true) Π_1 sentence ψ, such that $Q + \psi \not\leq T$ and $Q + \neg\psi \not\leq T$.

Proof. $\{\varphi: Q + \varphi \leq T\}$ is r.e. (Lemma 6.5) and monoconsistent with Q. Now use Lemma 2.1. ■

Our next result, Gödel's second incompleteness theorem, is a generalization of Theorem 2.4 (a). Since each t is primitive recursive, in PA we may use t as a function symbol.

Theorem 3. Suppose t: PA $\leq$ T.

(a) $T \nvdash t(Con_T)$.

(b) If $\tau(x)$ is any Σ_1 numeration of T, then $T \nvdash t(Con_\tau)$.

Proof. We prove (a); the proof of (b) is almost the same. Let φ be as in the proof of Theorem 1. Then $PA \vdash \neg\varphi \rightarrow Pr_T(t(\varphi))$. Moreover, $\neg\varphi$ being Σ_1, $PA \vdash \neg\varphi \rightarrow Pr_Q(\neg\varphi)$. Since Q is finite, it follows, by Fact 12 (Chapter 6), that $PA \vdash Pr_Q(\neg\varphi) \rightarrow Pr_T(\neg t(\varphi))$. We may now conclude that $PA \vdash \neg\varphi \rightarrow \neg Con_T$ and so $T \vdash t(Con_T) \rightarrow t(\varphi)$. But then, by Theorem 1, $T \nvdash t(Con_T)$, as desired. ■

We shall say that t is a *reflexive* interpretation of S in T, t: S $\leq^r$ T, if t: S $\leq$ T and for every k, $T \vdash t(Con_{T \restriction k})$. S is *reflexively interpretable,* S $\leq^r$ T, if there is a t such that t: S $\leq^r$ T. t is an *essentially reflexive* interpretation of S in T, t: S $\leq^{er}$ T, if t: S $\leq^r$ T′ for every pure extension T′ of T. S is *essentially reflexively interpretable,* S $\leq^{er}$ T, if there is a t such that t: S $\leq^{er}$ T.

As in Chapter 2, Theorem 3 has the following:

Corollary 1. If PA $\leq^r$ T, then T is not finitely axiomatizable.

$L_{ST} = \{\in\}$ is the language of (first order) set theory. ZF is Zermelo–Fraenkel set theory. Let t_s be the standard interpretation of arithmetic in set theory.

Fact 13. t_s: PA $\leq^{er}$ ZF.

Combining this with Corollary 1, we get (compare Corollary 2.1):

Corollary 2. No consistent pure extension of ZF is finitely axiomatizable.

This result will be strengthened in §§ 2 and 3 (Corollaries 3 and 7).

A nonreflexive interpretation of PA in ZF can be defined as follows. The theory $t_s^{-1}(ZF)$ is Σ_1–sound. Hence, by Corollary 6.9 (b), there is a faithful interpretation t′: PA $\trianglelefteq t_s^{-1}(ZF)$. Let $t = t_s t'$. Since t_s: $t_s^{-1}(ZF) \trianglelefteq$ ZF (Lemma 1 (b)), it follows that t: PA $\trianglelefteq$ ZF. There is a finite subtheory ZF$\restriction$k of ZF such that $PA \vdash Con_{ZF \restriction k} \rightarrow Con_{PA}$. (This follows from the fact that t_s is a "natural" interpretation of PA in a finite subtheory of ZF.) Since $PA \nvdash Con_{PA}$, it follows that $PA \nvdash Con_{ZF \restriction k}$. Since t is faithful, this implies that $ZF \nvdash t(Con_{ZF \restriction k})$ and so t is not reflexive.

§2. Axiomatizations. In this § we shall restrict ourselves to generalizing Theorem 4.2. We shall need the following generalization of part (a) of the fixed point lemma; the proof is left to the reader.

Lemma 2. Suppose t: PA ≤ T and let $\nu_n(x) := t(x = n)$. Let $\gamma(x)$ be any formula of L_T. There is then a sentence φ such that

$$T \vdash \varphi \leftrightarrow \exists y(\nu_\varphi(y) \wedge \gamma(y)).$$

We assume given a hierarchy $H = (H_0, H_1, \ldots)$ of the formulas of L_T satisfying closure conditions similar to those satisfied by the hierarchy $(\Sigma_0, \Sigma_1, \ldots)$. Thus, each H_k is a primitive recursive set of formulas, $H_k \subseteq H_{k+1}$, and $\bigcup\{H_k\colon k \in N\}$ is the set of all formulas of T. Let $H_k(x)$ be a PR binumeration of H_k.

We assume that for each k, there is an H_k *partial truth–definition* for H_k in T i.e. an H_k formula $Tr_k(x)$ such that for every H_k sentence φ,

(Tr_k) $T \vdash \varphi \leftrightarrow \exists x(\nu_\varphi(x) \wedge Tr_k(x))$.

We also assume that the formulas $Tr_k(x)$ are *mutatis mutandis* as in Fact 10 (a).

A set X of sentences of T is said to be H–*bounded* if there is a k such that $X \subseteq H_k$. Let

$$RFN^t_S = \{\forall x(t(H_k(x)) \wedge t(Pr_S(x)) \rightarrow Tr_k(x))\colon k \in N\}.$$

Theorem 4. Suppose t: PA ≤ T, $t(\Sigma_1) \subseteq H_0$, and $T \vdash RFN^t_S$. If X is any H–bounded set of sentences such that $T \dashv S + X$, then S + X is inconsistent.

Proof. This proof is essentially the same as the proof of Theorem 4.2. Let n be such that $X \subseteq H_n$. Let ψ be such that

(1) $T \vdash \psi \leftrightarrow \exists y\big(\nu_\psi(y) \wedge \forall xz(t(H_n(x)) \wedge Tr_n(x) \wedge t(z = (x \rightarrow y)) \rightarrow \neg t(Pr_S(z)))\big)$.

By assumption, we have

(2) $T \vdash \forall y\big(\nu_\psi(y) \rightarrow \forall xz(t(H_n(x)) \wedge t(z = (x \rightarrow y)) \wedge t(Pr_S(z)) \rightarrow (Tr_n(x) \rightarrow \psi))\big)$.

(1) and (2) imply that

(3) $T \vdash \psi$.

Suppose $T \dashv S + X$. By (3), there is then a conjunction θ of members of X such that $S + \theta \vdash \psi$. It follows that

(4) $T \vdash \exists z(\nu_{\theta \rightarrow \psi}(z) \wedge t(Pr_S(z)))$.

Also, by (1) and (3),

$$S + X \vdash \neg\exists z(\nu_{\theta \rightarrow \psi}(z) \wedge t(Pr_S(z))).$$

But then, by (4), S + X is inconsistent. ■

Theorem 4 can be applied to set theory. We define Σ^{ST}_k and Π^{ST}_k as follows. Let $\Sigma^{ST}_0 = \Pi^{ST}_0$ be the set of formulas of L_{ST} all of whose quantifiers are *bounded*, i.e. of the form $\exists x \in y$ or $\forall x \in y$. Σ^{ST}_{k+1} and Π^{ST}_{k+1} are then the least sets closed under bounded quantification such that $\Sigma^{ST}_k \subseteq \Pi^{ST}_{k+1}$, and $\Pi^{ST}_k \subseteq \Sigma^{ST}_{k+1}$, Σ^{ST}_{k+1} is closed under existential quantification and Π^{ST}_{k+1} is closed under universal quantification. A set X of sentences of L_{ST} is *bounded* if $X \subseteq \Sigma^{ST}_k$ for some k. We then have the following:

Fact 14. The assumptions of Theorem 4 are satisfied when $t = t_s$, $H_k = \Sigma^{ST}_{k+1}$, $T = ZF$, and $S = \emptyset$.

From Theorem 4 and Fact 14, we get:

Corollary 3. There is no bounded and consistent set X of sentences of L_{ST} such that $ZF \dashv X$.

§3. Interpretability. In this § we show that the relevant results of Chapter 6 generalize quite easily to the present more general setting. In using results from Chapter 6 we shall take advantage of the fact that in these results the theories S, S_0, etc. need not be formalized in L_A.

Theorem 5. If t: $PA \leq^r T$, then $T \leq t^{-1}(T)$ and so $t^{-1}(T) \equiv T$.

Proof. By assumption, $T \vdash t(Con_{T|k})$ for every k. It follows that $t^{-1}(T) \vdash Con_{T|k}$ for every k. But then, by Lemma 6.2, $T \leq t^{-1}(T)$. ■

Let us say that t: $PA \leq T$ is *optimal with respect to* Γ *sentences* if for every t′: $PA \leq T$ and every Γ sentence φ, if $T \vdash t'(\varphi)$, then $T \vdash t(\varphi)$.

Suppose t: $PA \leq T$. There is then a Σ_1 (true Π_2) sentence φ such that $t^{-1}(T) + \varphi \leq t^{-1}(T)$ and $t^{-1}(T) \nvdash \varphi$ (cf. Theorem 6.9 and the proof of Theorem 6.10). Let t′: $t^{-1}(T) + \varphi \leq t^{-1}(T)$ and set $t'' = tt'$. Then t″: $PA + \varphi \leq T$. Since $T \nvdash t(\varphi)$, it follows that t is not optimal with respect to Σ_1 (true Π_2) sentences. (If $t^{-1}(T)$ and φ are true, we can also achieve that $t''^{-1}(T)$ is true, since, by Corollary 6.9 (b), $t^{-1}(T) + \varphi \trianglelefteq t^{-1}(T)$.) In contrast to this we have the following:

Corollary 4. If t: $PA \leq^r T$, then t is optimal with respect to Π_1 sentences.

Proof. Suppose t′: $PA \leq T$. $t'^{-1}(T) \leq T$ and, by Theorem 5, $T \leq t^{-1}(T)$. Thus, $t'^{-1}(T) \leq t^{-1}(T)$. But then, by Theorem 6.6, $t'^{-1}(T) \dashv_{\Pi_1} t^{-1}(T)$. ■

Since $PA + Con_{PA} \leq ZF$, the nonreflexive t: $PA \leq ZF$ defined at the end of § 1 is not optimal with respect to Π_1 sentences.

From Fact 13 (this chapter) and Corollary 4 we get:

Corollary 5. If T is a pure extension of ZF, then t_s: $PA \leq T$ is optimal with respect to Π_1 sentences.

Corollary 5 can also be proved directly in the following way. Let T be as assumed. For any t: $PA \leq T$ and any model **M** of T, let $\mathbf{M}^t$ be the model of PA defined in **M** by t. In $\mathbf{M}^{ts}$ induction holds for every formula of L_{ST}. It follows that if t: $PA \leq T$, then $\mathbf{M}^{ts}$ is isomorphic to an initial segment of $\mathbf{M}^t$ (compare the proof of Theorem 6.7) and so every Π_1 sentence true in $\mathbf{M}^t$ is true in $\mathbf{M}^{ts}$.

From Lemma 1 (b) and Theorem 6.1, we get:

Lemma 3. There is a Σ_1 numeration $\tau'(x)$ of $t^{-1}(T)$ such that PA$\vdash$ $Con_T \rightarrow Con_{\tau'}$.

Theorem 6.2 can now be generalized as follows:

Theorem 6. Suppose t: PA $\leq^r$ T. Then T + $t(Con_T) \not\leq$ T.

Proof. Suppose T + $t(Con_T) \leq$ T. Then, by Lemma 1 (b) and (c), and Theorem 5, $t^{-1}(T) + Con_T \leq T \leq t^{-1}(T)$. Let $\tau'(x)$ be as in Lemma 3. It follows that $t^{-1}(T) + Con_{\tau'} \leq t^{-1}(T)$, contradicting Theorem 6.2. ■

Corollary 6. If T is a pure extension of ZF, then T + $t_s(Con_T) \not\leq$ T.

Theorem 6.3 has the following generalization:

Theorem 7. Suppose PA $\leq^r$ T. Then T is not interpretable in any finite subtheory of T.

Proof. Suppose $T \leq T \restriction m$. Let t be such that t: PA $\leq^r$ T. By Theorem 6.1, there is a Σ_1 numeration $\tau(x)$ of T such that PA$\vdash$ $Con_{T \restriction m} \rightarrow Con_\tau$. It follows that T$\vdash$ $t(Con_{T \restriction m}) \rightarrow t(Con_\tau)$. Also, by assumption, T$\vdash$ $t(Con_{T \restriction m})$. But then T$\vdash$ $t(Con_\tau)$, contradicting Theorem 3 (b). ■

Corollary 7. If T is a pure extension of ZF, then T is not interpretable in any finite subtheory of T.

The Orey–Hájek lemma in the present setting reads as follows:

Lemma 4. Suppose t: PA $\leq^r$ T. Then $S \leq T$ iff T$\vdash$ $t(Con_{S \restriction k})$ for every k.

Proof. By Lemma 6.2, $S \leq t^{-1}(T)$ iff $t^{-1}(T)\vdash$ $Con_{S \restriction k}$ for every k. ■

As in Chapter 6 we get, from Lemma 4, the following version of Orey's compactness theorem.

Theorem 8. Suppose PA $\leq^r$ T. Then $S \leq T$ iff for every k, $S \restriction k \leq T$.

Theorems 5 and 6.6 imply the following generalization of Theorem 6.6; we use A, B for pure extensions of T.

Theorem 9. Suppose t: PA $\leq^{er}$ T. Then $A \leq B$ iff $t^{-1}(A) \leq t^{-1}(B)$ iff $t^{-1}(A) \dashv_{\Pi_1} t^{-1}(B)$.

We conclude by generalizing Theorems 6.8 and 6.9; the generalization of Theorem 6.10 is left to the reader.

Theorem 10. If t: PA $\leq^{er}$ T, then T + $\neg$t(Con_T) $\leq$ T.

Proof. Let $\tau'(x)$ be as in Lemma 3. By Theorem 6.8, $t^{-1}(T) + \neg Con_{\tau'} \leq t^{-1}(T)$. It follows that $t^{-1}(T) + \neg Con_T \leq t^{-1}(T)$. But then, by Lemma 1 (c) and Theorem 5, T + $\neg t(Con_T) \leq t^{-1}(T + \neg t(Con_T)) \leq t^{-1}(T) + \neg Con_T \leq t^{-1}(T) \leq T$ and so T + $\neg t(Con_T) \leq T$, as desired. ■

Theorem 11. Suppose t: PA $\leq^{er}$ T and X is r.e. and monoconsistent with T. There is then a sentence φ such that T + $\varphi \leq$ T and $\varphi \notin X$; φ can be taken to be of the form t(ψ), where ψ is Σ_1.

Proof. Let Y = {θ: t(θ)$\in$X}. Then Y is r.e. and monoconsistent with $t^{-1}(T)$. By Theorem 6.9, there is a Σ_1 sentence ψ such that $t^{-1}(T) + \psi \leq t^{-1}(T)$ and $\psi \notin Y$. By Lemma 1 (c) and Theorem 5, T + t(ψ) $\leq$ T. Clearly, t(ψ)$\notin$X. ■

Theorem 11 has the following application, where GB is Gödel–Bernays (finite) set theory (compare Corollary 6.6).

Corollary 8. There is a Σ_1 sentence φ such that ZF + $t_s(\varphi) \leq$ ZF and GB + $t_s(\varphi) \not\leq$ GB.

Notes for Chapter 8.

Theorems 1 and 3 are, of course, (essentially) due to Gödel (1931), (1934) (cf. also Feferman (1960)). Theorem 2 is due to Montague (1957), (1962) (compare Exercise 6.1). For a definition of the standard interpretation t_s of arithmetic (theory of finite ordinals) in set theory, see, for example, Mendelson (1987) or Cohen (1966). Fact 13 and Corollary 2 are due to Montague (1961). Fact 14 is due to Lévy (1965). Theorem 4 and Corollary 3 are due to Kreisel and Lévy (1968), improving earlier work of Montague (1961). Corollaries 2 and 3 are given here as examples of applications of the corresponding general results; for a detailed discussion of similar, more general, and stronger results, see Kreisel and Lévy (1968). Lemma 4, Theorems 6, 7, 8, 9, 10, 11, and Corollary 8 are straightforward generalizations of the corresponding results in Chapter 6. The question if there is a sentence φ such that GB + $\varphi \leq$ GB and ZF + $\varphi \not\leq$ ZF, raised in Hájek (1971), was answered affirmatively by Solovay; for this and related results, cf. Hájek and Pudlák (1993).