A note on derivability conditions

Kurahashi, Taishi

(Citation)
The Journal of Symbolic Logic,85(3):1224-1253

(Issue Date)
2020-09-07

(Resource Type)
journal article

(Version)
Accepted Manuscript

A note on derivability conditions

Taishi Kurahashi

Abstract

We investigate relationships between versions of derivability conditions for provability predicates. We show several implications and non-implications between the conditions, and we discuss unprovability of consistency statements induced by derivability conditions. First, we classify already known versions of the second incompleteness theorem, and exhibit some new sets of conditions which are sufficient for unprovability of Hilbert–Bernays' consistency statement. Secondly, we improve Buchholz's schematic proof of provable Σ_1-completeness. Then among other things, we show that Hilbert–Bernays' conditions and Löb's conditions are mutually incomparable. We also show that neither Hilbert–Bernays' conditions nor Löb's conditions accomplish Gödel's original statement of the second incompleteness theorem.

1 Introduction

In his famous paper [8], Gödel proved the second incompleteness theorem with only a sketched proof. Gödel explained that by formalizing his proof of the first incompleteness theorem, the consistency statement $\exists x(\mathsf{Fml}(x) \land \neg \mathrm{Pr}_T(x))$ saying "there exists a T-unprovable formula" cannot be proved in T if T is consistent. To carry out his idea, it is desirable that the formula $\mathrm{Pr}_T(x)$ enjoys some natural properties as a formalization of the notion of T-provability. He wrote that a detailed proof would be presented in a forthcoming work, but such a paper was not published after all.

The first detailed proof of the second incompleteness theorem was presented in the second volume of *Grundlagen der Mathematik* [10] by Hilbert and Bernays. Especially they formulated a set of conditions for provability predicates which is sufficient for the second incompleteness theorem. Let $\mathrm{Pr}_T(x)$ be some Σ_1 provability predicate of T. They proved that if $\mathrm{Pr}_T(x)$ satisfies the following conditions **HB1**, **HB2** and **HB3**[1], then the consistency statement $\forall x(\mathsf{Fml}(x) \land \mathrm{Pr}_T(x) \to \neg \mathrm{Pr}_T(\dot{\neg} x))$ cannot be proved in T if T is consistent.

HB1 If $T \vdash \varphi \to \psi$, then $T \vdash \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \to \mathrm{Pr}_T(\ulcorner \psi \urcorner)$.

[1] More precisely, Hilbert–Bernays' conditions were originally stated on proof predicate $\mathfrak{B}(x, y)$ rather than on provability predicate $\mathrm{Pr}_T(x)$. For instance, the original statement of **HB1** is: If a formula with the number j is derived from a formula with the number i, then $\exists x \mathfrak{B}(x, i) \to \exists x \mathfrak{B}(x, j)$ is provable.

HB2 $T \vdash \mathrm{Pr}_T(\ulcorner \neg\varphi(x) \urcorner) \to \mathrm{Pr}_T(\ulcorner \neg\varphi(\dot{x}) \urcorner)$.

HB3 $T \vdash f(x) = 0 \to \mathrm{Pr}_T(\ulcorner f(\dot{x}) = 0 \urcorner)$ for every primitive recursive term $f(x)$.

Here $\ulcorner \varphi(\dot{x}) \urcorner$ is a primitive recursive term corresponding to a function calculating the Gödel number of the formula $\varphi(\overline{n})$ from n, where $\overline{n}$ is the numeral for n. These conditions are called the *Hilbert–Bernays derivability conditions*.

Löb [18] proved that if $\mathrm{Pr}_T(x)$ satisfies the following conditions **D1**, **D2** and **D3**, then Löb's theorem holds, that is, for any formula φ, if $T \vdash \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \to \varphi$, then $T \vdash \varphi$.

D1 If $T \vdash \varphi$, then $T \vdash \mathrm{Pr}_T(\ulcorner \varphi \urcorner)$.

D2 $T \vdash \mathrm{Pr}_T(\ulcorner \varphi \to \psi \urcorner) \to (\mathrm{Pr}_T(\ulcorner \varphi \urcorner) \to \mathrm{Pr}_T(\ulcorner \psi \urcorner))$.

D3 $T \vdash \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \to \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \urcorner)$.

Note that every provability predicate automatically satisfies **D1**. The conditions **D1** and **D2** were established by Hilbert and Bernays, and the condition **D3** was introduced by Löb. The conditions **D1**, **D2** and **D3** are nowadays called the *Hilbert–Bernays–Löb derivability conditions* which are well-known as sufficient conditions for a proof of the second incompleteness theorem. In fact, if T is consistent, then the unprovability of the consistency statement $\neg\mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner)$ in T is an immediate corollary of Löb's theorem. The Hilbert–Bernays–Löb derivability conditions together with Löb's theorem are basis for modal logical investigations of provability predicates (see [2, 5, 12, 22]).

Other sufficient conditions for the second incompleteness theorem were formulated by authors such as Jeroslow, Montagna and Buchholz. Jeroslow [13] proved that the following condition which is a variant of **D3** implies the unprovability of $\forall x(\mathsf{Fml}(x) \land \mathrm{Pr}_T(x) \to \neg\mathrm{Pr}_T(\dot{\neg} x))$.

- $T \vdash \mathrm{Pr}_T(t) \to \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(t) \urcorner)$ for every primitive recursive term t.

Notice that **D3** and Jeroslow's condition are instances of the following provable Σ_1-completeness because $\mathrm{Pr}_T(x)$ is Σ_1.

$\mathbf{\Sigma_1 C}$ If φ is a Σ_1 sentence, then $T \vdash \varphi \to \mathrm{Pr}_T(\ulcorner \varphi \urcorner)$.

Montagna [19] proved that the following two conditions are sufficient for Löb's theorem.

- $T \vdash \forall x($“x is a logical axiom” $\to \mathrm{Pr}_T(x))$.
- $T \vdash \forall x \forall y(\mathsf{Fml}(x) \land \mathsf{Fml}(y) \to (\mathrm{Pr}_T(x \dot{\to} y) \to (\mathrm{Pr}_T(x) \to \mathrm{Pr}_T(y))))$.

By Montagna's argument, we can conclude that these two conditions imply the unprovability of $\exists x(\mathsf{Fml}(x) \land \neg\mathrm{Pr}_T(x))$.

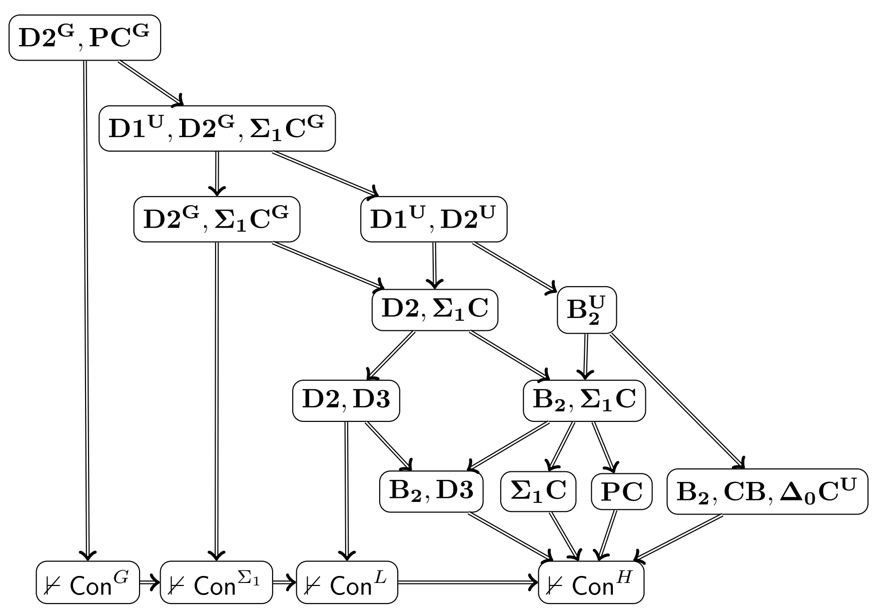
At last, in Buchholz's lecture note [6], the following condition was introduced and it was proved that this condition implies **D2** and $\mathbf{\Sigma_1 C}$.

- For all $m \geq 1$,
  if $T \vdash \forall \vec{x}(\varphi_1(\vec{x}) \to (\varphi_2(\vec{x}) \to (\cdots \to (\varphi_{m-1}(\vec{x}) \to \varphi_m(\vec{x}))\cdots)))$,
  then $T \vdash \forall \vec{x}(\mathrm{Pr}_T(\ulcorner \varphi_1(\dot{\vec{x}})\urcorner) \to (\mathrm{Pr}_T(\ulcorner \varphi_2(\dot{\vec{x}})\urcorner) \to$
  $(\cdots \to (\mathrm{Pr}_T(\ulcorner \varphi_{m-1}(\dot{\vec{x}})\urcorner) \to \mathrm{Pr}_T(\ulcorner \varphi_m(\dot{\vec{x}})\urcorner))\cdots)))$.

Thus Buchholz's condition implies the unprovability of $\neg\mathrm{Pr}_T(\ulcorner 0 \neq 0\urcorner)$.

Roughly speaking, every set of derivability conditions introduced above is sufficient for unprovability of consistency statements, but such a rough understanding does not allow us to grasp the situation of the second incompleteness theorem accurately. Strictly speaking, these sets of sufficient conditions do not induce the same consequence because there are three different consistency statements $\mathsf{Con}^H \equiv \forall x(\mathsf{Fml}(x) \land \mathrm{Pr}_T(x) \to \neg\mathrm{Pr}_T(\dot{\neg} x))$, $\mathsf{Con}^L \equiv \neg\mathrm{Pr}_T(\ulcorner 0 \neq 0\urcorner)$ and $\mathsf{Con}^G \equiv \exists x(\mathsf{Fml}(x) \land \neg\mathrm{Pr}_T(x))$ in our context, and each of these sets of conditions implies the unprovability of one of these consistency statements. Here superscripts 'H', 'L' and 'G' stand for Hilbert–Bernays, Löb and Gödel, respectively. It is easy to see that Con^H implies Con^L, and Con^L implies Con^G. However the converse implications do not hold in general.

In order to clarify the situation of several versions of derivability conditions, in this paper, we investigate relationships between the conditions. The following figure shows the situation for implications between prominent sets of conditions for Σ_1 formulas satisfying **D1**.



In Section 2, we introduce and investigate versions of derivability conditions. Each of these conditions is classified as one of three versions of derivability conditions, namely, local version, uniform version and global version. Among other things, we show that each of two new sets $\{\mathbf{D1}, \mathbf{B_2}, \mathbf{D3}\}$ and $\{\mathbf{D1}, \mathbf{PC}\}$ of derivability conditions is sufficient for the unprovability of the consistency statement Con^H (see the next section for precise definitions of these conditions). Then currently we know that four sets $\{\mathbf{B_2}, \mathbf{CB}, \mathbf{\Delta_0C^U}\}$, $\{\mathbf{D1}, \mathbf{B_2}, \mathbf{D3}\}$, $\{\mathbf{D1}, \mathbf{\Sigma_1C}\}$ and $\{\mathbf{D1}, \mathbf{PC}\}$ are sufficient for $T \nvdash \mathsf{Con}^H$, the set $\{\mathbf{D1}, \mathbf{D2}, \mathbf{D3}\}$ (Löb's conditions) is sufficient for $T \nvdash \mathsf{Con}^L$, and the set $\{\mathbf{D1}, \mathbf{D2^G}, \mathbf{PC^G}\}$ is sufficient for $T \nvdash \mathsf{Con}^G$. Here $\{\mathbf{B_2}, \mathbf{CB}, \mathbf{\Delta_0C^U}\}$, $\{\mathbf{D1}, \mathbf{\Sigma_1C}\}$ and $\{\mathbf{D1}, \mathbf{D2^G}, \mathbf{PC^G}\}$ correspond to Hilbert and Bernays' conditions, Jeroslow's conditions and Montagna's conditions, respectively.

In Section 3, we improve Buchholz's proof of provable Σ_1-completeness $\mathbf{\Sigma_1C}$. More precisely, we prove that if $\mathrm{Pr}_T(x)$ satisfies the following condition $\mathbf{B_2^U}$ which is precisely the $m = 2$ case of Buchholz's condition, then the uniform version of $\mathbf{\Sigma_1C}$ holds.

$\mathbf{B_2^U}$ If $T \vdash \forall \vec{x}\, (\varphi(\vec{x}) \to \psi(\vec{x}))$, then $T \vdash \forall \vec{x}(\mathrm{Pr}_T(\ulcorner \varphi(\dot{\vec{x}}) \urcorner) \to \mathrm{Pr}_T(\ulcorner \psi(\dot{\vec{x}}) \urcorner))$.

In Section 4, we give some examples of formulas, and from these examples, several non-implications between conditions are obtained. For instance, from our examples, we obtain that $\{\mathbf{B_2}, \mathbf{CB}, \mathbf{\Delta_0C^U}\}$, $\{\mathbf{D1}, \mathbf{B_2}, \mathbf{D3}\}$, $\{\mathbf{D1}, \mathbf{\Sigma_1C}\}$ and $\{\mathbf{D1}, \mathbf{PC}\}$ are pairwise incomparable, and each of them is not sufficient for $T \nvdash \mathsf{Con}^L$. Also we obtain that $\{\mathbf{D1}, \mathbf{D2}, \mathbf{D3}\}$ is not comparable with each of $\{\mathbf{B_2}, \mathbf{CB}, \mathbf{\Delta_0C^U}\}$, $\{\mathbf{D1}, \mathbf{\Sigma_1C}\}$ and $\{\mathbf{D1}, \mathbf{PC}\}$, and it is not sufficient for $T \nvdash \mathsf{Con}^G$. Furthermore, we show that even stronger set $\{\mathbf{D1^U}, \mathbf{D2^G}, \mathbf{\Sigma_1C^G}\}$ is not sufficient for $T \nvdash \mathsf{Con}^G$. From the last observation, we can say that both of the Hilbert–Bernays derivability conditions and the Hilbert–Bernays–Löb derivability conditions do not accomplish Gödel's original statement of the second incompleteness theorem.

2 Derivability conditions

Throughout this paper, S and T denote recursively axiomatized consistent extensions of Peano Arithmetic PA in the language of first-order arithmetic. The theory S is intended as a metatheory, and we assume that T is an extension of S. Let $\mathcal{L}_A$ be the language of arithmetic including $\{0, \mathsf{s}, +, \times\}$, and we can freely use terms corresponding to some primitive recursive functions. The numeral $\overline{n}$ for a natural number n is the closed term $\underbrace{\mathsf{s}(\mathsf{s}(\cdots \mathsf{s}}_{n \text{ times}}(0) \cdots))$. This explicit form of numerals is used in Section 3. We fix some natural Gödel numbering, and for each $\mathcal{L}_A$-formula φ, let $\ulcorner \varphi \urcorner$ be the numeral for the Gödel number of φ. Let $x \dot{\to} y$ and $\dot{\neg} x$ denote primitive recursive terms such that for any formulas φ and ψ, $\mathsf{PA} \vdash \ulcorner \varphi \urcorner \dot{\to} \ulcorner \psi \urcorner = \ulcorner \varphi \to \psi \urcorner$ and $\mathsf{PA} \vdash \dot{\neg} \ulcorner \varphi \urcorner = \ulcorner \neg \varphi \urcorner$.

Let $\Delta_0 = \Sigma_0 = \Pi_0$ be the set of all formulas whose quantifiers are all bounded. Let Σ_{n+1} and Π_{n+1} ($n \geq 0$) be the least sets of formulas satisfying the following conditions:

1. $\Sigma_n \cup \Pi_n \subseteq \Sigma_{n+1} \cap \Pi_{n+1}$;

2. Σ_{n+1} (resp. Π_{n+1}) is closed under conjunction, disjunction, bounded quantification, and existential (resp. universal) quantification;

3. If φ is in Σ_{n+1} (resp. Π_{n+1}), then $\neg\varphi$ is in Π_{n+1} (resp. Σ_{n+1});

4. If φ is in Σ_{n+1} (resp. Π_{n+1}) and ψ is in Π_{n+1} (resp. Σ_{n+1}), then $\varphi \to \psi$ is in Π_{n+1} (resp. Σ_{n+1}).

Throughout this paper, Γ denotes Σ_n or Π_n for some $n \geq 0$. We say a formula φ is Γ if $\varphi \in \Gamma$. A formula φ is said to be Δ_1 if it is provably equivalent to both some Σ_1 formula and some Π_1 formula in PA. Let $\mathsf{Fml}(x)$, $\mathsf{Sent}(x)$ and $\Sigma_z(x)$ be Δ_1 formulas saying that "x is the Gödel number of an $\mathcal{L}_A$-formula", "x is the Gödel number of an $\mathcal{L}_A$-sentence" and "x is the Gödel number of a Σ_z formula", respectively. We assume that PA can derive natural facts about these formulas such as $\forall z \exists x > z \mathsf{Fml}(x)$.

We say a formula $\mathrm{Pr}(x)$ is a *provability predicate* of a theory U (in PA) if it weakly represents the set of all theorems of U in PA, that is, for any natural number n, $\mathsf{PA} \vdash \mathrm{Pr}(\overline{n})$ if and only if n is the Gödel number of some theorem of U. Also we say a formula $\tau(v)$ is a *numeration* of U (in PA) if it weakly represents the set of all axioms of U in PA, that is, for any natural number n, $\mathsf{PA} \vdash \tau(\overline{n})$ if and only if n is the Gödel number of some axiom of U. For each numeration $\tau(v)$ of U, we can naturally construct a formula $\mathrm{Prf}_\tau(x, y)$ saying that "y is the code of a proof of a formula with the Gödel number x from the set of all sentences satisfying $\tau(v)$" (see Feferman [7]). We may assume $\mathsf{PA} \vdash \forall x \forall y (\mathrm{Prf}_\tau(x, y) \to x \leq y)$. If $\tau(v)$ is a Σ_n numeration of U for $n > 0$, then the formula $\mathrm{Pr}_\tau(x) :\equiv \exists y \mathrm{Prf}_\tau(x, y)$ is a Σ_n provability predicate of U. If it is not necessary to specify a particular numeration of U, $\mathrm{Prf}_U(x, y)$ and $\mathrm{Pr}_U(x)$ denote $\mathrm{Prf}_\tau(x, y)$ and $\mathrm{Pr}_\tau(x)$ for some fixed numeration $\tau(v)$ of U, respectively.

For each finitely axiomatized theory T_0, let $[T_0](x)$ be the formula $\bigvee_{\varphi \in T_0}(x = \ulcorner \varphi \urcorner)$. Then $[T_0](x)$ is a numeration of T_0. Let $\bigwedge T_0$ be the conjunction of all axioms of T_0, and let $\mathrm{Pr}_\emptyset(x)$ be a natural provability predicate of first-order predicate calculus in the language $\mathcal{L}_A$. Then the following lemma holds (see Feferman [7]).

Lemma 2.1 (Formalized deduction theorem)**.** *For any finitely axiomatized theory T_0, $\mathsf{PA} \vdash \forall x (\mathrm{Pr}_{[T_0]}(x) \leftrightarrow \mathrm{Pr}_\emptyset(\ulcorner \bigwedge T_0 \urcorner \dot{\to} x))$.*

Throughout this paper, the formula $\Phi(x)$ is intended to denote some provability predicate of T. However, we deal with more general situations, that is, $\Phi(x)$ may not be any provability predicate of T. In this section, we introduce a lot of conditions for $\Phi(x)$ which are satisfied by naturally constructed provability predicates $\mathrm{Pr}_T(x)$. The remainder of this section is separated into three subsections, and in each of these subsections, we introduce local derivability conditions, uniform derivability conditions and global derivability conditions, respectively.

For each formula $\Phi(x)$, we define four kinds of consistency statements based on $\Phi(x)$.

Definition 2.2.

1. $\mathsf{Con}^H_\Phi :\equiv \forall x(\mathsf{Fml}(x) \land \Phi(x) \to \neg\Phi(\dot{\neg} x))$.
2. $\mathsf{Con}^L_\Phi :\equiv \neg\Phi(\ulcorner 0 \neq 0 \urcorner)$.
3. $\mathsf{Con}^G_\Phi :\equiv \exists x(\mathsf{Fml}(x) \land \neg\Phi(x))$.
4. $\mathsf{Con}^{\Sigma_1}_\Phi :\equiv \exists x(\Sigma_1(x) \land \mathsf{Sent}(x) \land \neg\Phi(x))$.

The first consistency statement Con^H_Φ is adopted in Hilbert and Bernays [10] and Feferman [7]. The second sentence Con^L_Φ is the most tractable one, and it is widely used in the context of modal logical investigations of provability predicates. Gödel [8] stated his second incompleteness theorem with the consistency statement Con^G_Φ. The last consistency statement $\mathsf{Con}^{\Sigma_1}_\Phi$ states that there exists a T-unprovable Σ_1 sentence.

2.1 Local derivability conditions

We introduce the weakest version of derivability conditions which are called local derivability conditions.

Definition 2.3 (Local derivability conditions)**.**

D1 If $T \vdash \varphi$, then $S \vdash \Phi(\ulcorner \varphi \urcorner)$ for any formula φ.

D2 $S \vdash \Phi(\ulcorner \varphi \to \psi \urcorner) \to (\Phi(\ulcorner \varphi \urcorner) \to \Phi(\ulcorner \psi \urcorner))$ for any formulas φ and ψ.

D3 $S \vdash \Phi(\ulcorner \varphi \urcorner) \to \Phi(\ulcorner \Phi(\ulcorner \varphi \urcorner) \urcorner)$ for any formula φ.

ΓC $S \vdash \varphi \to \Phi(\ulcorner \varphi \urcorner)$ for any Γ sentence φ.

$\mathbf{B_m}$ ($m \geq 1$) If $T \vdash \bigwedge_{0<i<m} \varphi_i \to \varphi_m$, then $S \vdash \bigwedge_{0<i<m} \Phi(\ulcorner \varphi_i \urcorner) \to \Phi(\ulcorner \varphi_m \urcorner)$ for any formulas $\varphi_1, \ldots, \varphi_m$.

PC $S \vdash \mathrm{Pr}_\emptyset(\ulcorner \varphi \urcorner) \to \Phi(\ulcorner \varphi \urcorner)$ for any formula φ.

The condition **D1** is automatically satisfied by all provability predicates of T. The conditions **D2**, **D3** and $\mathbf{\Sigma_1 C}$ were introduced by Hilbert and Bernays [10], Löb [18] and Feferman [7], respectively. It is known that natural provability predicates $\mathrm{Pr}_T(x)$ satisfy full local derivability conditions. In particular, Feferman proved $\mathbf{\Sigma_1 C}$ for the provability predicate $\mathrm{Pr}_{\mathsf{Q}}(x)$ of Robinson's arithmetic Q (cf. [23]). The conditions $\mathbf{B_m}$ ($m \geq 1$) were introduced by Buchholz [6]. The condition $\mathbf{B_1}$ is precisely **D1**, and the condition $\mathbf{B_2}$ is precisely the condition **HB1** described in the introduction. The condition $\mathbf{B_2}$ was also discussed by Montagna [19] and Visser [24]. The last condition **PC** says that $\Phi(x)$ contains predicate calculus.

We prove the basic implications between local derivability conditions. For example, the first clause of the following proposition says that if a formula $\Phi(x)$ satisfies **D1**, then it also satisfies $\mathbf{\Delta_0 C}$.

Proposition 2.4.

1. *$\mathbf{D1} \Rightarrow \mathbf{\Delta_0 C}$.*
2. *$\mathbf{\Delta_0 C}$ and $\mathbf{B_m}$ for some $m \geq 1 \Rightarrow \mathbf{D1}$.*
3. *$\mathbf{B_3} \Rightarrow \mathbf{D2}$.*
4. *The following are equivalent:*
   (a) *$\mathbf{D1}$ and $\mathbf{D2}$.*
   (b) *$\mathbf{B_m}$ for all $m \geq 1$.*
   (c) *$\mathbf{D1}$ and $\mathbf{B_m}$ for some $m \geq 3$.*
   (d) *$\mathbf{\Delta_0 C}$ and $\mathbf{B_m}$ for some $m \geq 3$.*
5. *If $\Phi(x)$ is a Γ formula, then $\mathbf{\Gamma C} \Rightarrow \mathbf{D3}$.*
6. *$\mathbf{B_2}$ and $\mathbf{PC} \iff \mathbf{B_2}$ and $\mathbf{\Sigma_1 C}$.*
7. *$\mathbf{B_2}$ and $\mathbf{PC} \Rightarrow \mathbf{D1}$.*
8. *$\mathbf{D1}$, $\mathbf{D2}$ and $\mathbf{PC} \iff \mathbf{D1}$, $\mathbf{D2}$ and $\mathbf{\Sigma_1 C}$.*

Proof. 1. Suppose $\Phi(x)$ satisfies $\mathbf{D1}$. Let φ be any Δ_0 sentence. Then φ is decidable in PA. If $\mathsf{PA} \vdash \varphi$, then $S \vdash \Phi(\ulcorner \varphi \urcorner)$ by $\mathbf{D1}$, and hence $S \vdash \varphi \to \Phi(\ulcorner \varphi \urcorner)$. If $\mathsf{PA} \vdash \neg\varphi$, then $S \vdash \varphi \to \Phi(\ulcorner \varphi \urcorner)$.

2. Suppose $\Phi(x)$ satisfies $\mathbf{\Delta_0 C}$ and $\mathbf{B_m}$ for some $m \geq 1$. Let φ be any formula with $T \vdash \varphi$. Then $T \vdash \underbrace{0 = 0 \land \cdots \land 0 = 0}_{m-1} \to \varphi$. By $\mathbf{B_m}$, we have $S \vdash \Phi(\ulcorner 0 = 0 \urcorner) \to \Phi(\ulcorner \varphi \urcorner)$. By $\mathbf{\Delta_0 C}$, $S \vdash 0 = 0 \to \Phi(\ulcorner 0 = 0 \urcorner)$, and hence $S \vdash \Phi(\ulcorner 0 = 0 \urcorner)$. We conclude $S \vdash \Phi(\ulcorner \varphi \urcorner)$.

3. Since $T \vdash (\varphi \to \psi) \land \varphi \to \psi$, we obtain $S \vdash \Phi(\ulcorner \varphi \to \psi \urcorner) \land \Phi(\ulcorner \varphi \urcorner) \to \Phi(\ulcorner \psi \urcorner)$ by $\mathbf{B_3}$.

4. $(a) \Rightarrow (b)$ is well-known in the context of modal logic. $(b) \Rightarrow (c)$ is trivial. $(c) \Leftrightarrow (d)$ follows from clauses 1 and 2. We prove $(c) \Rightarrow (a)$: Suppose $\Phi(x)$ satisfies $\mathbf{D1}$ and $\mathbf{B_m}$ for some $m \geq 3$. By clause 3, it suffices to prove that $\Phi(x)$ satisfies $\mathbf{B_3}$. Suppose $T \vdash \varphi_1 \land \varphi_2 \to \varphi_3$. Then $T \vdash \varphi_1 \land \varphi_2 \land \underbrace{0 = 0 \land \cdots \land 0 = 0}_{m-3} \to \varphi_3$. By $\mathbf{B_m}$, we obtain $S \vdash \Phi(\ulcorner \varphi_1 \urcorner) \land \Phi(\ulcorner \varphi_2 \urcorner) \land \Phi(\ulcorner 0 = 0 \urcorner) \to \Phi(\ulcorner \varphi_3 \urcorner)$. By $\mathbf{D1}$, we have $S \vdash \Phi(\ulcorner 0 = 0 \urcorner)$. Hence $S \vdash \Phi(\ulcorner \varphi_1 \urcorner) \land \Phi(\ulcorner \varphi_2 \urcorner) \to \Phi(\ulcorner \varphi_3 \urcorner)$.

5. Trivial.

6. $(\Rightarrow)$: Assume that $\Phi(x)$ satisfies $\mathbf{B_2}$ and $\mathbf{PC}$. Let φ be any Σ_1 sentence. Let T_0 be some finite subtheory of T containing Robinson's arithmetic Q. By $\mathbf{PC}$, $S \vdash \mathrm{Pr}_\emptyset(\ulcorner \bigwedge T_0 \to \varphi \urcorner) \to \Phi(\ulcorner \bigwedge T_0 \to \varphi \urcorner)$. Here $\mathrm{Pr}_\emptyset(\ulcorner \bigwedge T_0 \to \varphi \urcorner)$ is equivalent to $\mathrm{Pr}_{[T_0]}(\ulcorner \varphi \urcorner)$ by formalized deduction theorem (Lemma 2.1), and therefore we obtain $S \vdash \mathrm{Pr}_{[T_0]}(\ulcorner \varphi \urcorner) \to \Phi(\ulcorner \bigwedge T_0 \to \varphi \urcorner)$. Since T_0 is a subtheory of T, we have $T \vdash (\bigwedge T_0 \to \varphi) \to \varphi$. By $\mathbf{B_2}$, $S \vdash \Phi(\ulcorner \bigwedge T_0 \to \varphi \urcorner) \to \Phi(\ulcorner \varphi \urcorner)$.

Thus we obtain $S \vdash \mathrm{Pr}_{[T_0]}(\ulcorner \varphi \urcorner) \to \Phi(\ulcorner \varphi \urcorner)$. Since T_0 contains Q, $\mathbf{\Sigma_1 C}$ holds for $\mathrm{Pr}_{[T_0]}(x)$, and hence $S \vdash \varphi \to \mathrm{Pr}_{[T_0]}(\ulcorner \varphi \urcorner)$. Therefore $S \vdash \varphi \to \Phi(\ulcorner \varphi \urcorner)$.

($\Leftarrow$): Suppose $\Phi(x)$ satisfies $\mathbf{B_2}$ and $\mathbf{\Sigma_1 C}$. Let φ be any formula. Since $\mathrm{Pr}_{\emptyset}(\ulcorner \varphi \urcorner)$ is a Σ_1 sentence, $S \vdash \mathrm{Pr}_{\emptyset}(\ulcorner \varphi \urcorner) \to \Phi(\ulcorner \mathrm{Pr}_{\emptyset}(\ulcorner \varphi \urcorner) \urcorner)$. Since T is an extension of PA, $T \vdash \mathrm{Pr}_{\emptyset}(\ulcorner \varphi \urcorner) \to \varphi$ by the reflexiveness of PA (see [17]). By $\mathbf{B_2}$, $S \vdash \Phi(\ulcorner \mathrm{Pr}_{\emptyset}(\ulcorner \varphi \urcorner) \urcorner) \to \Phi(\ulcorner \varphi \urcorner)$. Therefore $S \vdash \mathrm{Pr}_{\emptyset}(\ulcorner \varphi \urcorner) \to \Phi(\ulcorner \varphi \urcorner)$.

7. This follows from clauses 2 and 6.

8. This equivalence follows from clauses 4 and 6. □

Before describing several versions of the second incompleteness theorem, we prepare two propositions.

Proposition 2.5.

1. *If $\Phi(x)$ satisfies* **D1**, *then* $S \vdash \mathsf{Con}^H_\Phi \to \mathsf{Con}^L_\Phi$.
2. $\mathsf{PA} \vdash \mathsf{Con}^L_\Phi \to \mathsf{Con}^{\Sigma_1}_\Phi$.
3. $\mathsf{PA} \vdash \mathsf{Con}^{\Sigma_1}_\Phi \to \mathsf{Con}^G_\Phi$.

Proof. 1. Suppose $\Phi(x)$ satisfies **D1**, then $S \vdash \Phi(\ulcorner 0 = 0 \urcorner)$. Since $\mathsf{PA} \vdash \mathsf{Con}^H_\Phi \to (\Phi(\ulcorner 0 = 0 \urcorner) \to \neg \Phi(\ulcorner 0 \neq 0 \urcorner))$, we have $S \vdash \mathsf{Con}^H_\Phi \to \mathsf{Con}^L_\Phi$.

Clauses 2 and 3 are obvious. □

The following proposition is a part of Gödel's first incompleteness theorem.

Proposition 2.6. *Let φ be a sentence satisfying $\mathsf{PA} \vdash \varphi \leftrightarrow \neg \Phi(\ulcorner \varphi \urcorner)$. If $\Phi(x)$ satisfies* **D1**, *then $T \nvdash \varphi$.*

Proof. Suppose $\Phi(x)$ satisfies **D1**. If $T \vdash \varphi$, then by **D1**, $S \vdash \Phi(\ulcorner \varphi \urcorner)$. By the choice of φ, $S \vdash \neg \varphi$. This contradicts the consistency of T because T is an extension of S. Therefore $T \nvdash \varphi$. □

It is well-known that for proofs of the second incompleteness theorem, the Hilbert–Bernays–Löb derivability conditions **D1**, **D2** and **D3** are sufficient. This is essentially due to Löb (see [5, 17]).

Theorem 2.7 (Löb [18])**.** *If $\Phi(x)$ satisfies* **D1**, **D2** *and* **D3**, *then* $T \nvdash \mathsf{Con}^L_\Phi$.

Notice that $\{\mathbf{D1}, \mathbf{B_2}, \mathbf{D3}\}$ is weaker than $\{\mathbf{D1}, \mathbf{D2}, \mathbf{D3}\}$ by Proposition 2.4.4. For the former conditions, we obtain another version of the second incompleteness theorem.

Theorem 2.8. *If $\Phi(x)$ satisfies* **D1**, $\mathbf{B_2}$ *and* **D3**, *then* $T \nvdash \mathsf{Con}^H_\Phi$.

Proof. Suppose $\Phi(x)$ satisfies **D1**, $\mathbf{B_2}$ and **D3**. Let φ be a sentence satisfying $\mathsf{PA} \vdash \varphi \leftrightarrow \neg \Phi(\ulcorner \varphi \urcorner)$. The existence of such a sentence φ follows from the Fixed Point Lemma (see [17]). Since $T \vdash \Phi(\ulcorner \varphi \urcorner) \to \neg \varphi$, we have $S \vdash \Phi(\ulcorner \Phi(\ulcorner \varphi \urcorner) \urcorner) \to \Phi(\ulcorner \neg \varphi \urcorner)$ by $\mathbf{B_2}$. By **D3**, $S \vdash \Phi(\ulcorner \varphi \urcorner) \to \Phi(\ulcorner \Phi(\ulcorner \varphi \urcorner) \urcorner)$. Thus $S \vdash \Phi(\ulcorner \varphi \urcorner) \to \Phi(\ulcorner \neg \varphi \urcorner)$, and hence $S \vdash \neg \varphi \to \exists x (\mathsf{Fml}(x) \land \Phi(x) \land \Phi(\dot{\neg} x))$. It follows $S \vdash \mathsf{Con}^H_\Phi \to \varphi$. By Proposition 2.6, $T \nvdash \varphi$, and thus $T \nvdash \mathsf{Con}^H_\Phi$. □

Jeroslow [13] proved that if $\mathcal{L}_A$ contains sufficiently many primitive recursive terms and if $\Phi(x)$ satisfies **D1** and $S \vdash \Phi(t) \to \Phi(\ulcorner \Phi(t) \urcorner)$ for all primitive recursive terms t, then $T \nvdash \mathsf{Con}^H_\Phi$. That is to say, in Theorem 2.8, if we strengthen the condition **D3** in this way, then the condition $\mathbf{B_2}$ can be omitted. As a consequence, Jeroslow remarked that if $\Phi(x)$ is a Γ formula, then the conditions **D1** and $\mathbf{\Gamma C}$ are sufficient for the unprovability of Con^H_Φ in Jersolow's setting of language. We show that this is also the case without using such sufficiently many primitive recursive terms.

Theorem 2.9 (Jeroslow [13]; Kreisel and Takeuti [15])**.** *If $\Phi(x)$ is a Γ formula satisfying* **D1** *and* $\mathbf{\Gamma C}$*, then $T \nvdash \mathsf{Con}^H_\Phi$.*

Proof. Let φ be a Γ sentence such that $\mathsf{PA} \vdash \varphi \leftrightarrow \Phi(\ulcorner \neg\varphi \urcorner)$. By Proposition 2.6, $T \nvdash \neg\varphi$ because of **D1**. By $\mathbf{\Gamma C}$ and the choice of φ, $S \vdash \varphi \to \Phi(\ulcorner \varphi \urcorner) \land \Phi(\ulcorner \neg\varphi \urcorner)$. Then we have $S \vdash \varphi \to \neg\mathsf{Con}^H_\Phi$. Therefore $T \nvdash \mathsf{Con}^H_\Phi$. $\square$

By Proposition 2.4.8 and Theorem 2.7, if $\Phi(x)$ is a Σ_1 formula satisfying **D1**, **D2** and **PC**, then $T \nvdash \mathsf{Con}^L_\Phi$. Also by Proposition 2.4.6 and Theorem 2.9, if $\Phi(x)$ is a Σ_1 formula satisfying **D1**, $\mathbf{B_2}$ and **PC**, then $T \nvdash \mathsf{Con}^H_\Phi$. We improve the latter statement as follows.

Theorem 2.10. *If $\Phi(x)$ is a Σ_1 formula satisfying* **D1** *and* **PC***, then $T \nvdash \mathsf{Con}^H_\Phi$.*

Proof. Suppose that $\Phi(x)$ is Σ_1 and satisfies **D1** and **PC**. Let T_0 be a finite subtheory of T containing Q. Let φ be a Σ_1 sentence satisfying $\mathsf{PA} \vdash \varphi \leftrightarrow \Phi(\ulcorner \neg(\bigwedge T_0 \to \varphi) \urcorner)$. By **PC** and formalized deduction theorem, we have $S \vdash \mathrm{Pr}_{[T_0]}(\ulcorner \varphi \urcorner) \to \Phi(\ulcorner \bigwedge T_0 \to \varphi \urcorner)$. By $\mathbf{\Sigma_1 C}$ for $\mathrm{Pr}_{[T_0]}(x)$, $S \vdash \varphi \to \Phi(\ulcorner \bigwedge T_0 \to \varphi \urcorner)$. Since $\mathsf{PA} \vdash \varphi \to \Phi(\ulcorner \neg(\bigwedge T_0 \to \varphi) \urcorner)$ by the choice of φ, we obtain $S \vdash \varphi \to \neg\mathsf{Con}^H_\Phi$.

If $T \vdash \mathsf{Con}^H_\Phi$, then $T \vdash \neg\varphi$. Also $T \vdash \bigwedge T_0 \land \neg\varphi$, and this means $T \vdash \neg(\bigwedge T_0 \to \varphi)$. By **D1**, $S \vdash \Phi(\ulcorner \neg(\bigwedge T_0 \to \varphi) \urcorner)$, and hence $S \vdash \varphi$. This is a contradiction. Therefore $T \nvdash \mathsf{Con}^H_\Phi$. $\square$

Remark 2.11. The following makeshift condition $\mathbf{\Sigma_1 C^-}$ is of course weaker than $\mathbf{\Sigma_1 C}$ if $\bigwedge \emptyset \to \varphi$ is identical to φ.

$\mathbf{\Sigma_1 C^-}$ There exists a finite subtheory T_0 of T such that for any Σ_1 sentence φ, $S \vdash \varphi \to \Phi(\ulcorner \bigwedge T_0 \to \varphi \urcorner)$.

Our proof of Proposition 2.4.6 ($\Rightarrow$) actually shows two implications "$\mathbf{PC} \Rightarrow \mathbf{\Sigma_1 C^-}$" and "$\{\mathbf{B_2}, \mathbf{\Sigma_1 C^-}\} \Rightarrow \mathbf{\Sigma_1 C}$". Also our proof of Theorem 2.10 essentially shows that if $\Phi(x)$ is a Σ_1 formula satisfying **D1** and $\mathbf{\Sigma_1 C^-}$, then $T \nvdash \mathsf{Con}^H_\Phi$. Then Theorem 2.9 in the case $\Gamma = \Sigma_1$ and Theorem 2.10 directly follow from these observations.

In this section, we have seen that $\{\mathbf{D1}, \mathbf{D2}, \mathbf{D3}\}$ is sufficient for $T \nvdash \mathsf{Con}^L_\Phi$ (Theorem 2.7), and $\{\mathbf{D1}, \mathbf{B_2}, \mathbf{D3}\}$ is sufficient for $T \nvdash \mathsf{Con}^H_\Phi$ (Theorem 2.8). Also for Σ_1 formulas $\Phi(x)$, each of $\{\mathbf{D1}, \mathbf{\Sigma_1 C}\}$ and $\{\mathbf{D1}, \mathbf{PC}\}$ is sufficient for $T \nvdash \mathsf{Con}^H_\Phi$ (Theorems 2.9 and 2.10). From examples of formulas given in Section

4, the following non-implications are obtained. These non-implications show that these unprovability results are optimal. For example, the third clause in the following list means that there exists a Σ_1 formula $\Phi(x)$ satisfying both **D1** and **D2** such that $T \vdash \mathsf{Con}^H_\Phi$.

- $\{\mathbf{D1}, \mathbf{D2}, \mathbf{\Sigma_1 C}\} \not\Rightarrow T \nvdash \mathsf{Con}^H_\Phi$ (Fact 4.3).
- $\{\Phi \in \Sigma_1, \mathbf{D2}, \mathbf{D3}, \mathbf{\Sigma_1 C}, \mathbf{PC}\} \not\Rightarrow T \nvdash \mathsf{Con}^H_\Phi$ (Proposition 4.1).
- $\{\Phi \in \Sigma_1, \mathbf{D1}, \mathbf{D2}\} \not\Rightarrow T \nvdash \mathsf{Con}^H_\Phi$ (Fact 4.5.1).
- $\{\Phi \in \Sigma_1, \mathbf{D1}, \mathbf{D3}\} \not\Rightarrow T \nvdash \mathsf{Con}^H_\Phi$ (Fact 4.5.2).
- $\{\Phi \in \Sigma_1, \mathbf{D1}, \mathbf{B_2}, \mathbf{D3}\} \not\Rightarrow T \nvdash \mathsf{Con}^L_\Phi$ (Fact 4.6.3).
- $\{\Phi \in \Sigma_1, \mathbf{D1}, \mathbf{\Sigma_1 C}, \mathbf{PC}\} \not\Rightarrow T \nvdash \mathsf{Con}^L_\Phi$ (Proposition 4.4).
- $\{\Phi \in \Sigma_1, \mathbf{D1}, \mathbf{D2}, \mathbf{\Sigma_1 C}\} \not\Rightarrow T \nvdash \mathsf{Con}^{\Sigma_1}_\Phi$ (Proposition 4.10).

These non-implications show that none of $\{\mathbf{D1}, \mathbf{B_2}, \mathbf{D3}\}$, $\{\mathbf{D1}, \mathbf{\Sigma_1 C}\}$ and $\{\mathbf{D1}, \mathbf{PC}\}$ implies $\{\mathbf{D1}, \mathbf{D2}, \mathbf{D3}\}$. Moreover we obtain the following non-implications.

- $\{\Phi \in \Sigma_1, \mathbf{D1}, \mathbf{D2}, \mathbf{D3}\} \not\Rightarrow \mathbf{\Sigma_1 C}$ (Proposition 4.12). By Proposition 2.4.6, this is equivalent to $\{\Phi \in \Sigma_1, \mathbf{D1}, \mathbf{D2}, \mathbf{D3}\} \not\Rightarrow \mathbf{PC}$.
- $\{\Phi \in \Sigma_1, \mathbf{D1}, \mathbf{\Sigma_1 C}, \mathbf{PC}\} \not\Rightarrow \mathbf{B_2}$ (Proposition 4.4).
- $\{\Phi \in \Sigma_1, \mathbf{D1}, \mathbf{\Sigma_1 C}\} \not\Rightarrow \mathbf{PC}$ (Proposition 4.13).
- $\{\Phi \in \Sigma_1, \mathbf{D1}, \mathbf{PC}\} \not\Rightarrow \mathbf{\Sigma_1 C}$ (Proposition 4.14).

Consequently, $\{\mathbf{D1}, \mathbf{B_2}, \mathbf{D3}\}$, $\{\mathbf{D1}, \mathbf{\Sigma_1 C}\}$ and $\{\mathbf{D1}, \mathbf{PC}\}$ are pairwise incomparable. Also $\{\mathbf{D1}, \mathbf{D2}, \mathbf{D3}\}$ is incomparable with each of $\{\mathbf{D1}, \mathbf{\Sigma_1 C}\}$ and $\{\mathbf{D1}, \mathbf{PC}\}$.

2.2 Uniform derivability conditions

In this subsection, we introduce and investigate uniform derivability conditions. Let $\varphi(\vec{x})$ be an abbreviation for $\varphi(x_0, \ldots, x_k)$ for some k.

Definition 2.12 (Uniform derivability conditions)**.**

$\mathbf{D1^U}$ If $T \vdash \forall \vec{x}\, \varphi(\vec{x})$, then $S \vdash \forall \vec{x}\, \Phi(\ulcorner \varphi(\dot{\vec{x}}) \urcorner)$ for any formula $\varphi(\vec{x})$.

$\mathbf{D2^U}$ $S \vdash \forall \vec{x}\, (\Phi(\ulcorner \varphi(\dot{\vec{x}}) \to \psi(\dot{\vec{x}}) \urcorner) \to (\Phi(\ulcorner \varphi(\dot{\vec{x}}) \urcorner) \to \Phi(\ulcorner \psi(\dot{\vec{x}}) \urcorner)))$ for any formulas $\varphi(\vec{x})$ and $\psi(\vec{x})$.

$\mathbf{D3^U}$ $S \vdash \forall \vec{x}\, (\Phi(\ulcorner \varphi(\dot{\vec{x}}) \urcorner) \to \Phi(\ulcorner \Phi(\ulcorner \varphi(\dot{\vec{x}}) \urcorner) \urcorner))$ for any formula $\varphi(\vec{x})$.

$\mathbf{\Gamma C^U}$ $S \vdash \forall \vec{x}\, (\varphi(\vec{x}) \to \Phi(\ulcorner \varphi(\dot{\vec{x}}) \urcorner))$ for any Γ formula $\varphi(\vec{x})$.

$\mathbf{B^U_m}$ $(m \geq 1)$ If $T \vdash \forall \vec{x}\left(\bigwedge_{0<i<m} \varphi_i(\vec{x}) \to \varphi_m(\vec{x})\right)$,

then $S \vdash \forall \vec{x}\left(\bigwedge_{0<i<m} \Phi(\ulcorner \varphi_i(\dot{\vec{x}})\urcorner) \to \Phi(\ulcorner \varphi_m(\dot{\vec{x}})\urcorner)\right)$

for any formulas $\varphi_1(\vec{x}), \ldots, \varphi_m(\vec{x})$.

CB $S \vdash \Phi(\ulcorner \forall \vec{x}\, \varphi(\vec{x})\urcorner) \to \forall \vec{x}\, \Phi(\ulcorner \varphi(\dot{\vec{x}})\urcorner)$ for any formula $\varphi(\vec{x})$.

$\mathbf{PC^U}$ $S \vdash \forall \vec{x}(\mathrm{Pr}_\emptyset(\ulcorner \varphi(\dot{\vec{x}})\urcorner) \to \Phi(\ulcorner \varphi(\dot{\vec{x}})\urcorner))$ for any formula $\varphi(\vec{x})$.

Usual proofs of the Hilbert–Bernays–Löb derivability conditions **D1**, **D2** and **D3** (in books such as [5]) are demonstrated by showing stronger uniform derivability conditions $\mathbf{D1^U}$, $\mathbf{D2^U}$ and $\mathbf{\Sigma_1C^U}$. Notice that the natural provability predicates $\mathrm{Pr}_T(x)$ satisfy full uniform derivability conditions.

As in the local version, the conditions $\mathbf{B^U_m}$ $(m \geq 1)$ were introduced by Buchholz [6], and $\mathbf{B^U_1}$ is precisely $\mathbf{D1^U}$. The condition **CB** claims that sentences corresponding to the Converse Barcan Formula investigated in predicate modal logic (see [11]) are provable. Notice that the condition **HB2** described in the introduction seems to be a variant of the condition **CB**. It is easy to see that each of uniform derivability conditions is stronger than the corresponding local version. Moreover, uniform derivability conditions are strictly stronger than local derivability conditions (see Proposition 4.9 in Section 4).

As in the local version, we obtain the following proposition.

Proposition 2.13.

1. $\mathbf{\Delta_0C}$ *and* $\mathbf{B^U_m}$ *for some* $m \geq 1 \Rightarrow \mathbf{D1^U}$.
2. $\mathbf{B^U_3} \Rightarrow \mathbf{D2^U}$.
3. *The following are equivalent:*
   (a) $\mathbf{D1^U}$ *and* $\mathbf{D2^U}$.
   (b) $\mathbf{B^U_m}$ *for all* $m \geq 1$.
   (c) $\mathbf{D1^U}$ *and* $\mathbf{B^U_m}$ *for some* $m \geq 3$.
4. *If* $\Phi(x)$ *is a* Γ *formula, then* $\mathbf{\Gamma C^U} \Rightarrow \mathbf{D3^U}$.
5. $\mathbf{B^U_2}$ *and* $\mathbf{PC^U} \iff \mathbf{B^U_2}$ *and* $\mathbf{\Sigma_1C^U}$.
6. $\mathbf{B^U_2}$ *and* $\mathbf{PC^U} \Rightarrow \mathbf{D1^U}$.
7. $\mathbf{D1^U}$, $\mathbf{D2^U}$ *and* $\mathbf{PC^U} \iff \mathbf{D1^U}$, $\mathbf{D2^U}$ *and* $\mathbf{\Sigma_1C^U}$.

The condition **CB** is related to other conditions.

Proposition 2.14.

1. **D1** *and* **CB** $\Rightarrow \mathbf{D1^U}$.

2. $\mathbf{B_2^U} \Rightarrow \mathbf{CB}$.

3. $\mathbf{D2^U}$ *and* $\mathbf{PC^U} \Rightarrow \mathbf{CB}$.

4. The following are equivalent:

   (a) $\mathbf{D1^U}$ *and* $\mathbf{D2^U}$.

   (b) $\mathbf{D1}$, $\mathbf{B_2^U}$ *and* $\mathbf{D2^U}$.

   (c) $\mathbf{D1}$, $\mathbf{CB}$ *and* $\mathbf{D2^U}$.

Proof. 1. Suppose that $\Phi(x)$ satisfies **D1** and **CB**. Assume $T \vdash \forall \vec{x}\, \varphi(\vec{x})$. Then $S \vdash \Phi(\ulcorner \forall \vec{x}\, \varphi(\vec{x}) \urcorner)$ by **D1**. Since $S \vdash \Phi(\ulcorner \forall \vec{x}\, \varphi(\vec{x}) \urcorner) \to \forall \vec{x}\, \Phi(\ulcorner \varphi(\dot{\vec{x}}) \urcorner)$ by **CB**, we have $S \vdash \forall \vec{x}\, \Phi(\ulcorner \varphi(\dot{\vec{x}}) \urcorner)$.

2. Suppose that $\Phi(x)$ satisfies $\mathbf{B_2^U}$. Since $T \vdash \forall \vec{x}\, \varphi(\vec{x}) \to \varphi(\vec{x})$, we have $S \vdash \Phi(\ulcorner \forall \vec{x}\, \varphi(\vec{x}) \urcorner) \to \Phi(\ulcorner \varphi(\dot{\vec{x}}) \urcorner)$ by $\mathbf{B_2^U}$. Therefore $S \vdash \Phi(\ulcorner \forall \vec{x}\, \varphi(\vec{x}) \urcorner) \to \forall \vec{x}\, \Phi(\ulcorner \varphi(\dot{\vec{x}}) \urcorner)$.

3. Suppose $\Phi(x)$ satisfies $\mathbf{D2^U}$ and $\mathbf{PC^U}$. Let $\varphi(\vec{x})$ be any formula. Since $\forall \vec{x} \varphi(\vec{x}) \to \varphi(\vec{x})$ is provable in predicate calculus, $S \vdash \mathrm{Pr}_\emptyset(\ulcorner \forall \vec{x} \varphi(\vec{x}) \to \varphi(\dot{\vec{x}}) \urcorner)$ by $\mathbf{D1^U}$ for $\mathrm{Pr}_\emptyset(x)$. From $\mathbf{PC^U}$, $S \vdash \Phi(\ulcorner \forall \vec{x} \varphi(\vec{x}) \to \varphi(\dot{\vec{x}}) \urcorner)$. Then by $\mathbf{D2^U}$, $S \vdash \Phi(\ulcorner \forall \vec{x} \varphi(\vec{x}) \urcorner) \to \Phi(\ulcorner \varphi(\dot{\vec{x}}) \urcorner)$. Thus $S \vdash \Phi(\ulcorner \forall \vec{x}\, \varphi(\vec{x}) \urcorner) \to \forall \vec{x}\, \Phi(\ulcorner \varphi(\dot{\vec{x}}) \urcorner)$.

4. The implications $(a) \Rightarrow (b)$, $(b) \Rightarrow (c)$ and $(c) \Rightarrow (a)$ follow from Proposition 2.13.3, clause 2 and clause 1, respectively. □

The following corollary immediately follows from clauses 1, 2 and 3 of Proposition 2.14.

Corollary 2.15.

1. $\mathbf{D1}$ *and* $\mathbf{B_2^U} \Rightarrow \mathbf{D1^U}$.

2. $\mathbf{D1}$, $\mathbf{D2^U}$ *and* $\mathbf{PC^U} \Rightarrow \mathbf{D1^U}$.

Hilbert and Bernays [10] proved that if a Σ_1 formula $\Phi(x)$ satisfies the conditions **HB1**, **HB2** and **HB3** described in the introduction, then $T \nvdash \mathsf{Con}_\Phi^H$. In our framework, the Hilbert–Bernays derivability conditions can be replaced by the conditions $\mathbf{B_2}$, **CB** and $\mathbf{\Delta_0 C^U}$ without any substantial change. Then we obtain the following version of the second incompleteness theorem.

Theorem 2.16 (Hilbert and Bernays [10])**.** *If* $\Phi(x)$ *is a* Σ_1 *formula satisfying* $\mathbf{B_2}$, $\mathbf{CB}$ *and* $\mathbf{\Delta_0 C^U}$, *then* $T \nvdash \mathsf{Con}_\Phi^H$.

Proof. Suppose that $\Phi(x)$ is Σ_1 and satisfies $\mathbf{B_2}$, **CB** and $\mathbf{\Delta_0 C^U}$. Let φ be a Π_1 sentence satisfying $\mathsf{PA} \vdash \varphi \leftrightarrow \neg \Phi(\ulcorner \varphi \urcorner)$. Let $\delta(x)$ be a Δ_0 formula with $\mathsf{PA} \vdash \varphi \leftrightarrow \forall x \delta(x)$. Then by $\mathbf{B_2}$, $S \vdash \Phi(\ulcorner \varphi \urcorner) \to \Phi(\ulcorner \forall x \delta(x) \urcorner)$. By **CB**, we obtain

$$S \vdash \neg \varphi \to \forall x \Phi(\ulcorner \delta(\dot{x}) \urcorner). \tag{1}$$

On the other hand, $S \vdash \neg\delta(x) \to \Phi(\ulcorner \neg\delta(\dot{x}) \urcorner)$ by $\mathbf{\Delta_0 C^U}$. Then $S \vdash \exists x \neg\delta(x) \to \exists x \Phi(\ulcorner \neg\delta(\dot{x}) \urcorner)$. Hence $S \vdash \neg\varphi \to \exists x \Phi(\ulcorner \neg\delta(\dot{x}) \urcorner)$. By combining this with (1), we obtain

$$S \vdash \neg\varphi \to \exists x(\Phi(\ulcorner \delta(\dot{x}) \urcorner) \land \Phi(\ulcorner \neg\delta(\dot{x}) \urcorner)).$$

It follows $S \vdash \neg\varphi \to \exists x(\mathsf{Fml}(x) \land \Phi(x) \land \Phi(\dot{\neg} x))$, and hence $S \vdash \mathsf{Con}^H_\Phi \to \varphi$. By Proposition 2.4.2, $\Phi(x)$ satisfies $\mathbf{D1}$. Then by Proposition 2.6, $T \nvdash \varphi$. Therefore we conclude $T \nvdash \mathsf{Con}^H_\Phi$. $\square$

Theorem 2.16 is optimal in the sense of the following non-implications from Section 4.

- $\{\mathbf{D1}, \mathbf{B_2}, \mathbf{CB}, \mathbf{\Delta_0 C^U}\} \not\Rightarrow T \nvdash \mathsf{Con}^H_\Phi$ (Fact 4.3).
- $\{\Phi \in \Sigma_1, \mathbf{CB}, \mathbf{\Delta_0 C^U}\} \not\Rightarrow T \nvdash \mathsf{Con}^H_\Phi$ (Proposition 4.1).
- $\{\Phi \in \Sigma_1, \mathbf{B_2}, \mathbf{CB}\} \not\Rightarrow T \nvdash \mathsf{Con}^H_\Phi$ (Proposition 4.2).
- $\{\Phi \in \Sigma_1, \mathbf{D1}, \mathbf{B_2}, \mathbf{\Delta_0 C^U}\} \not\Rightarrow T \nvdash \mathsf{Con}^H_\Phi$ (Fact 4.6.1).
- $\{\Phi \in \Sigma_1, \mathbf{D1}, \mathbf{B_2}, \mathbf{CB}, \mathbf{\Delta_0 C^U}\} \not\Rightarrow T \nvdash \mathsf{Con}^L_\Phi$ (Fact 4.6.2).

Notice that $\{\mathbf{B_2}, \mathbf{CB}, \mathbf{\Delta_0 C^U}\}$ is equivalent to $\{\mathbf{D1}, \mathbf{B_2}, \mathbf{CB}, \mathbf{\Delta_0 C^U}\}$ by Proposition 2.4.2. For the latter condition, we do not know if $\{\Phi \in \Sigma_1, \mathbf{D1}, \mathbf{B_2}, \mathbf{CB}, \mathbf{\Delta_0 C^U}\}$ is optimal to conclude $T \nvdash \mathsf{Con}^H_\Phi$ or not.

Problem 2.17.

1. *Is there a Σ_1 provability predicate satisfying $\mathbf{D1}$, $\mathbf{CB}$ and $\mathbf{\Delta_0 C^U}$ such that $T \vdash \mathsf{Con}^H_\Phi$?*
2. *Is there a Σ_1 provability predicate satisfying $\mathbf{D1}$, $\mathbf{B_2}$ and $\mathbf{CB}$ such that $T \vdash \mathsf{Con}^H_\Phi$?*

The following two non-implications from Section 4 indicate that $\{\mathbf{B_2}, \mathbf{CB}, \mathbf{\Delta_0 C^U}\}$ is incomparable with each of $\{\mathbf{D1}, \mathbf{D2}, \mathbf{D3}\}$, $\{\mathbf{D1}, \mathbf{B_2}, \mathbf{D3}\}$, $\{\mathbf{D1}, \mathbf{\Sigma_1 C}\}$ and $\{\mathbf{D1}, \mathbf{PC}\}$.

- $\{\Phi \in \Sigma_1, \mathbf{B_2}, \mathbf{CB}, \mathbf{\Delta_0 C^U}\} \not\Rightarrow \mathbf{D3}$ (Fact 4.6.2).
- $\{\Phi \in \Sigma_1, \mathbf{D1}, \mathbf{D2}, \mathbf{\Sigma_1 C}\} \not\Rightarrow \mathbf{CB}$ (Proposition 4.9).

Usual proof of $\mathbf{\Sigma_1 C^U}$ (in books such as [5]) proceeds by induction on the construction of Σ_1 formulas, and it requires much effort. In the lecture note [6] by Buchholz, an elegant schematic proof of $\mathbf{\Sigma_1 C^U}$ is presented. More precisely, it is proved that for a proof of $\mathbf{\Sigma_1 C^U}$, the assumption "$\mathbf{B^U_m}$ for all $m \geq 1$" is sufficient. By Proposition 2.13.3, this assumption is equivalent to $\{\mathbf{D1^U}, \mathbf{D2^U}\}$. Hence Buchholz's work is stated as follows.

Theorem 2.18 (Buchholz [6]). $\mathbf{D1^U}$ *and* $\mathbf{D2^U} \Rightarrow \mathbf{\Sigma_1 C^U}$.

In Rautenberg's book [21], a schematic proof of $\mathbf{\Sigma_1 C^U}$ based on Buchholz's argument is presented. As a corollary to Theorem 2.18, we obtain the following version of the second incompleteness theorem.

Corollary 2.19. *If $\Phi(x)$ is a Σ_1 formula satisfying $\mathbf{D1^U}$ and $\mathbf{D2^U}$, then $T \nvdash \mathsf{Con}^L_\Phi$.*

Notice that $\{\mathbf{D1^U}, \mathbf{D2^U}\}$ implies $\{\mathbf{D1}, \mathbf{B_2^U}\}$ by Proposition 2.13.3. The following theorem improves Buchholz's Theorem 2.18 which will be proved in the next section.

Theorem 2.20. $\mathbf{D1}$ *and* $\mathbf{B_2^U} \Rightarrow \mathbf{\Sigma_1 C^U}$.

This theorem says that only the $m = 1, 2$ cases of Buchholz's assumption are sufficient to prove $\mathbf{\Sigma_1 C^U}$. We will also prove that Theorem 2.20 is actually an improvement of Theorem 2.18 (see Theorem 4.15 below). Interestingly, for Σ_1 formulas, $\{\mathbf{D1}, \mathbf{B_2^U}\}$ implies $\{\mathbf{D1}, \mathbf{B_2}, \mathbf{D3}\}$, $\{\mathbf{D1}, \mathbf{\Sigma_1 C}\}$, $\{\mathbf{D1}, \mathbf{PC}\}$ and $\{\mathbf{B_2}, \mathbf{CB}, \mathbf{\Delta_0 C^U}\}$ by Theorem 2.20 and Proposition 2.13, and each of them is sufficient for $T \nvdash \mathsf{Con}^H_\Phi$. As a consequence, we obtain the following corollary.

Corollary 2.21. *If $\Phi(x)$ is a Σ_1 formula satisfying $\mathbf{D1}$ and $\mathbf{B_2^U}$, then $T \nvdash \mathsf{Con}^H_\Phi$.*

Related to Corollary 2.21, we propose the following problem.

Problem 2.22. *Is there a Σ_1 formula $\Phi(x)$ satisfying $\mathbf{D1}$ and $\mathbf{B_2^U}$ such that $T \vdash \mathsf{Con}^L_\Phi$?*

In contrast to the consistency statements Con^H_Φ and Con^L_Φ, Proposition 4.10 in Section 4 shows that the full uniform derivability conditions are not sufficient for the unprovability of $\mathsf{Con}^{\Sigma_1}_\Phi$ and Con^G_Φ.

From Theorem 2.20 and Proposition 2.13.5, we obtain the following corollary.

Corollary 2.23. $\mathbf{D1}$ *and* $\mathbf{B_2^U} \Rightarrow \mathbf{PC^U}$.

Moreover, we show that $\mathbf{D1}$ and $\mathbf{B_2^U}$ imply a stronger version of $\mathbf{PC^U}$. For $n \geq 0$, let $\mathsf{True}_{\Sigma_n}(x)$ be a natural formula saying that "x is a true Σ_n sentence" (cf. Hájek and Pudlák [9]).

Proposition 2.24. *If $\Phi(x)$ satisfies $\mathbf{D1}$ and $\mathbf{B_2^U}$, then for $n \geq 0$,*

$$S \vdash \forall x(\Sigma_n(x) \land \mathrm{Pr}_\emptyset(x) \to \Phi(\ulcorner \mathsf{True}_{\Sigma_n}(\dot{x}) \urcorner)).$$

Proof. Suppose that $\Phi(x)$ satisfies $\mathbf{D1}$ and $\mathbf{B_2^U}$, and let $n \geq 0$. By Theorem 2.20, $\Phi(x)$ satisfies $\mathbf{\Sigma_1 C^U}$, and hence $S \vdash \Sigma_n(x) \land \mathrm{Pr}_\emptyset(x) \to \Phi(\ulcorner \Sigma_n(\dot{x}) \land \mathrm{Pr}_\emptyset(\dot{x}) \urcorner)$. By reflexiveness, $T \vdash \Sigma_n(x) \land \mathrm{Pr}_\emptyset(x) \to \mathsf{True}_{\Sigma_n}(x)$. Then $S \vdash \Phi(\ulcorner \Sigma_n(\dot{x}) \land \mathrm{Pr}_\emptyset(\dot{x}) \urcorner) \to \Phi(\ulcorner \mathsf{True}_{\Sigma_n}(\dot{x}) \urcorner)$ by $\mathbf{B_2^U}$. We conclude $S \vdash \Sigma_n(x) \land \mathrm{Pr}_\emptyset(x) \to \Phi(\ulcorner \mathsf{True}_{\Sigma_n}(\dot{x}) \urcorner)$. □

2.3 Global derivability conditions

At last, we introduce the strongest version of derivability conditions. They are called global derivability conditions.

Definition 2.25 (Global derivability conditions).

$\mathbf{D2^G}$ $S \vdash \forall x \forall y(\mathsf{Fml}(x) \land \mathsf{Fml}(y) \to (\Phi(x \dot{\to} y) \to (\Phi(x) \to \Phi(y))))$.

$\mathbf{D3^G}$ $S \vdash \forall x(\mathsf{Fml}(x) \to (\Phi(x) \to \Phi(\ulcorner \Phi(\dot{x}) \urcorner)))$.

$\mathbf{\Gamma C^G}$ $S \vdash \forall x(\mathsf{True}_\Gamma(x) \to \Phi(x))$.

$\mathbf{PC^G}$ $S \vdash \forall x(\mathsf{Fml}(x) \to (\mathrm{Pr}_\emptyset(x) \to \Phi(x)))$.

The condition $\mathbf{D2^G}$ for provability predicates $\mathrm{Pr}_T(x)$ was proved in Feferman [7]. Montagna [19] investigated the condition $\mathbf{D2^G}$. The condition $\mathbf{\Sigma_1 C^G}$ for $\mathrm{Pr}_{\mathsf{Q}}(x)$ is explicitly stated in the book [9]. Global derivability conditions are strictly stronger than uniform derivability conditions (see Proposition 4.10).

We can prove the following proposition as in the uniform version.

Proposition 2.26.

1. *If $\Phi(x)$ is a Γ formula, then $\mathbf{\Gamma C^U} \Rightarrow \mathbf{D3^G}$.*
2. *$\mathbf{D1}$, $\mathbf{D2^G}$ and $\mathbf{PC^G} \Rightarrow \mathbf{\Sigma_1 C^G}$.*

Proposition 2.26.2 was stated in von Bülow [26] and Visser [25].

Consistency statements are enhanced by global derivability conditions.

Proposition 2.27.

1. *If $\Phi(x)$ satisfies $\mathbf{D2^G}$ and $\mathbf{PC^G}$, then $S \vdash \mathsf{Con}^G_\Phi \to \mathsf{Con}^H_\Phi$.*
2. *If $\Phi(x)$ satisfies $\mathbf{D1}$, $\mathbf{D2^G}$ and $\mathbf{PC^G}$, then Con^H_Φ, Con^L_Φ and Con^G_Φ are mutually equivalent in S.*
3. *If $\Phi(x)$ satisfies $\mathbf{D2^G}$ and $\mathbf{\Sigma_1 C^G}$, then Con^L_Φ and $\mathsf{Con}^{\Sigma_1}_\Phi$ are equivalent in S.*

Proof. 1. Suppose $\Phi(x)$ satisfies $\mathbf{D2^G}$ and $\mathbf{PC^G}$. Since $\mathsf{PA} \vdash \forall x \forall y(\mathsf{Fml}(x) \land \mathsf{Fml}(y) \to \mathrm{Pr}_\emptyset(x \dot{\to} (\dot{\neg} x \dot{\to} y)))$, $S \vdash \forall x \forall y(\mathsf{Fml}(x) \land \mathsf{Fml}(y) \to \Phi(x \dot{\to} (\dot{\neg} x \dot{\to} y)))$ by $\mathbf{PC^G}$. Hence $\forall x \forall y(\mathsf{Fml}(x) \land \mathsf{Fml}(y) \land \Phi(x) \land \Phi(\dot{\neg} x) \to \Phi(y))$ is provable in S by $\mathbf{D2^G}$. This sentence is equivalent to $\mathsf{Con}^G_\Phi \to \mathsf{Con}^H_\Phi$.

2. This follows from Proposition 2.5 and clause 1.

3. Suppose $\Phi(x)$ satisfies $\mathbf{D2^G}$ and $\mathbf{\Sigma_1 C^G}$. By Proposition 2.5, it suffices to show $S \vdash \mathsf{Con}^{\Sigma_1}_\Phi \to \mathsf{Con}^L_\Phi$. Since $\mathsf{PA} \vdash \neg \mathsf{True}_{\Sigma_1}(\ulcorner 0 \neq 0 \urcorner)$, $\mathsf{PA} \vdash \Sigma_1(x) \land \mathsf{Sent}(x) \to \mathsf{True}_{\Sigma_1}(\ulcorner 0 \neq 0 \urcorner \dot{\to} x)$. By $\mathbf{\Sigma_1 C^G}$, $S \vdash \Sigma_1(x) \land \mathsf{Sent}(x) \to \Phi(\ulcorner 0 \neq 0 \urcorner \dot{\to} x)$. By $\mathbf{D2^G}$, $S \vdash \Sigma_1(x) \land \mathsf{Sent}(x) \to (\Phi(\ulcorner 0 \neq 0 \urcorner) \to \Phi(x))$. Thus $S \vdash \mathsf{Con}^{\Sigma_1}_\Phi \to \mathsf{Con}^L_\Phi$. ☐

From Theorems 2.7 and 2.10, and Proposition 2.27, we obtain the following corollary.

Corollary 2.28.

1. *If* $\Phi(x)$ *is a* Σ_1 *formula satisfying* $\mathbf{D1}$*,* $\mathbf{D2^G}$ *and* $\mathbf{PC^G}$*, then* $T \nvdash \mathsf{Con}^G_\Phi$*.*

2. *If* $\Phi(x)$ *is a* Σ_1 *formula satisfying* $\mathbf{D1}$*,* $\mathbf{D2^G}$ *and* $\mathbf{\Sigma_1 C^G}$*, then* $T \nvdash \mathsf{Con}^{\Sigma_1}_\Phi$*.*

Corollary 2.15.2 and Proposition 2.26.2 show that $\{\mathbf{D1^U}, \mathbf{D2^G}, \mathbf{\Sigma_1 C^G}\}$ is weaker than $\{\mathbf{D1}, \mathbf{D2^G}, \mathbf{PC^G}\}$. Moreover, Proposition 4.11 in Section 4 shows the following interesting non-implication:

- $\{\Phi \in \Sigma_1, \mathbf{D1^U}, \mathbf{D2^G}, \mathbf{\Sigma_1 C^G}\} \not\Rightarrow T \nvdash \mathsf{Con}^G_\Phi$.

Hence in contrast to local and uniform versions, $\{\mathbf{D1^U}, \mathbf{D2^G}, \mathbf{\Sigma_1 C^G}\}$ is strictly weaker than $\{\mathbf{D1}, \mathbf{D2^G}, \mathbf{PC^G}\}$. Also this non-implication indicates that global derivability conditions except for $\mathbf{PC^G}$ are not sufficient for the unprovability of Gödel's consistency statement Con^G_Φ even if Φ is Σ_1. This shows that neither Hilbert–Bernays' conditions nor Löb's conditions accomplish Gödel's original statement of the second incompleteness theorem.

Let $\mathsf{LogAx}(x)$ be a suitable Δ_1 formula representing the set of all logical axioms of predicate calculus formulated in Feferman's paper [7]. In Feferman's formulation, the sole inference rule is modus ponens, and the generalization rule is admissible (see Result 2.1 in [7]). The following condition was introduced by Montagna [19].

Definition 2.29.

$\mathbf{Ax}$ $S \vdash \forall x(\mathsf{LogAx}(x) \to \Phi(x))$.

The condition $\mathbf{Ax}$ is related to the condition $\mathbf{PC^G}$.

Proposition 2.30.

1. $\mathbf{PC^G} \Rightarrow \mathbf{Ax}$*.*

2. $\mathbf{D2^G}$ *and* $\mathbf{Ax} \Rightarrow \mathbf{PC^G}$*.*

3. *If* $\Phi(x)$ *satisfies* $\mathbf{D1}$*, then for any sentence* φ*,* $S \vdash \mathsf{LogAx}(\ulcorner \varphi \urcorner) \to \Phi(\ulcorner \varphi \urcorner)$*.*

Proof. 1. This is because $\mathsf{PA} \vdash \forall x(\mathsf{LogAx}(x) \to \mathrm{Pr}_\emptyset(x))$.

2. Let $\mathrm{Pr}'_\emptyset(x)$ be a natural provability predicate of the predicate calculus formulated in Feferman's framework. Then $\mathsf{PA} \vdash \forall x(\mathsf{Fml}(x) \to (\mathrm{Pr}_\emptyset(x) \to \mathrm{Pr}'_\emptyset(x)))$ holds by induction inside PA. Since S proves that $\Phi(x)$ contains axioms of $\mathrm{Pr}'_\emptyset(x)$ by $\mathbf{Ax}$ and that $\Phi(x)$ is closed under the inference rule of $\mathrm{Pr}'_\emptyset(x)$ by $\mathbf{D2^G}$, S proves $\forall x(\mathsf{Fml}(x) \to (\mathrm{Pr}'_\emptyset(x) \to \Phi(x)))$ by induction inside S. Hence $S \vdash \forall x(\mathsf{Fml}(x) \to (\mathrm{Pr}_\emptyset(x) \to \Phi(x)))$ holds.

3. Let φ be any sentence. If φ is a logical axiom, then $T \vdash \varphi$. By $\mathbf{D1}$, $S \vdash \Phi(\ulcorner \varphi \urcorner)$. If φ is not a logical axiom, then $S \vdash \neg \mathsf{LogAx}(\ulcorner \varphi \urcorner)$. In either case, we obtain $S \vdash \mathsf{LogAx}(\ulcorner \varphi \urcorner) \to \Phi(\ulcorner \varphi \urcorner)$. ☐

Montagna [19] proved that if $\Phi(x)$ satisfies $\mathbf{D1}$, $\mathbf{D2^G}$ and $\mathbf{Ax}$, then $\mathbf{D3}$ is redundant for a proof of Löb's theorem. From Propositions 2.26 and 2.30, and Corollaries 2.15.2 and 2.28, we obtain the following improvement of Montagna's result.

Corollary 2.31 (Montagna [19])**.**

1. *$\mathbf{D1}$, $\mathbf{D2^G}$ and $\mathbf{Ax} \Rightarrow \mathbf{D1^U}$ and $\mathbf{\Sigma_1 C^G}$.*
2. *If $\Phi(x)$ is a Σ_1 formula satisfying $\mathbf{D1}$, $\mathbf{D2^G}$ and $\mathbf{Ax}$, then $T \nvdash \mathsf{Con}_\Phi^G$.*

3 Proof of Theorem 2.20

In this section, we prove Theorem 2.20, that is, we prove that if $\Phi(x)$ satisfies $\mathbf{D1}$ and $\mathbf{B_2^U}$, then $\Phi(x)$ satisfies $\mathbf{\Sigma_1 C^U}$. Thus in the rest of this section, we fix a formula $\Phi(x)$ satisfying $\mathbf{D1}$ and $\mathbf{B_2^U}$. Then by Corollary 2.15.1, $\Phi(x)$ also satisfies $\mathbf{D1^U}$. First, we prove a lemma, that is an essential application of the condition $\mathbf{B_2^U}$.

Lemma 3.1. *Let $\varphi(\vec{x})$ and $\psi(\vec{x})$ be any formulas. If $S \vdash \varphi(\vec{x}) \to \Phi(\ulcorner \varphi(\dot{\vec{x}}) \urcorner)$ and $\mathsf{PA} \vdash \varphi(\vec{x}) \leftrightarrow \psi(\vec{x})$, then $S \vdash \psi(\vec{x}) \to \Phi(\ulcorner \psi(\dot{\vec{x}}) \urcorner)$.*

Proof. If $\mathsf{PA} \vdash \varphi(\vec{x}) \leftrightarrow \psi(\vec{x})$, then by $\mathbf{B_2^U}$, we have

$$S \vdash \Phi(\ulcorner \varphi(\dot{\vec{x}}) \urcorner) \leftrightarrow \Phi(\ulcorner \psi(\dot{\vec{x}}) \urcorner).$$

Then the lemma follows immediately. □

We may assume that every Σ_1 $\mathcal{L}_A$-formula is PA-provably equivalent to some Σ_1 formula written in the language $\{0, \mathsf{s}, +, \times\}$. Therefore, in proving Theorem 2.20, it suffices to show $S \vdash \sigma(\vec{x}) \to \Phi(\ulcorner \sigma(\dot{\vec{x}}) \urcorner)$ for any Σ_1 formula $\sigma(\vec{x})$ in the language $\{0, \mathsf{s}, +, \times\}$. Hence in the rest of this section, we assume that our terms and formulas are written in $\{0, \mathsf{s}, +, \times\}$. Before proving Theorem 2.20, we prepare several lemmas.

Lemma 3.2. *For any formula $\varphi(\vec{y}, v)$,*

$$\mathsf{PA} \vdash \ulcorner \varphi(\dot{\vec{y}}, \dot{v}) \urcorner [\mathsf{s}(x)/v] = \ulcorner \varphi(\dot{\vec{y}}, \mathsf{s}(\dot{x})) \urcorner,$$

where $\ulcorner \varphi(\dot{\vec{y}}, \dot{v}) \urcorner [\mathsf{s}(x)/v]$ is the result of substituting $\mathsf{s}(x)$ for v of $\ulcorner \varphi(\dot{\vec{y}}, \dot{v}) \urcorner$.

Proof. This is because our numeral $\overline{n}$ is defined by applying s to 0 n times. Then the lemma can be proved by induction on the constructions of terms and formulas. We give only an outline of a proof.

For example, we assume that our Gödel number $\mathsf{gn}(t)$ of a term t is defined so that $\mathsf{gn}(\mathsf{s}(t)) = \langle 0, \mathsf{gn}(t) \rangle$, where $\langle \cdot, \cdot \rangle$ is a primitive recursive paring function. Then we can define a primitive recursive function $\mathsf{num}(x)$ calculating $n \mapsto \mathsf{gn}(\overline{n})$ satisfying $\mathsf{num}(\mathsf{s}(x)) = \langle 0, \mathsf{num}(x) \rangle$. This is proved in PA and corresponds to $\ulcorner \dot{v} \urcorner [\mathsf{s}(x)/v] = \ulcorner \mathsf{s}(\dot{x}) \urcorner$. Then by using properties of $\ulcorner \cdot \urcorner$ such as $\mathsf{PA} \vdash \ulcorner \mathsf{s}(t) \urcorner = \langle 0, \ulcorner t \urcorner \rangle$, we can show $\mathsf{PA} \vdash \ulcorner t(\dot{\vec{y}}, \dot{v}) \urcorner [\mathsf{s}(x)/v] = \ulcorner t(\dot{\vec{y}}, \mathsf{s}(\dot{x})) \urcorner$ for any term $t(\vec{y}, v)$. Then we can prove the lemma by using properties of $\ulcorner \cdot \urcorner$. □

Lemma 3.3. *Let $\varphi(\vec{x}, v)$ be any formula. If $S \vdash \varphi(\vec{x}, v) \to \Phi(\ulcorner \varphi(\dot{\vec{x}}, \dot{v}) \urcorner)$, then $S \vdash \exists v \varphi(\vec{x}, v) \to \Phi(\ulcorner \exists v \varphi(\dot{\vec{x}}, v) \urcorner)$.*

Proof. Suppose $S \vdash \varphi(\vec{x}, v) \to \Phi(\ulcorner \varphi(\dot{\vec{x}}, \dot{v}) \urcorner)$. Since $T \vdash \varphi(\vec{x}, v) \to \exists v \varphi(\vec{x}, v)$, we have $S \vdash \Phi(\ulcorner \varphi(\dot{\vec{x}}, \dot{v}) \urcorner) \to \Phi(\ulcorner \exists v \varphi(\dot{\vec{x}}, v) \urcorner)$ by $\mathbf{B_2^U}$. Hence $S \vdash \varphi(\vec{x}, v) \to \Phi(\ulcorner \exists v \varphi(\dot{\vec{x}}, v) \urcorner)$. Therefore we conclude $S \vdash \exists v \varphi(\vec{x}, v) \to \Phi(\ulcorner \exists v \varphi(\dot{\vec{x}}, v) \urcorner)$. $\square$

Lemma 3.4. *For any natural number k and any variables $x_0, \ldots, x_k, z_0, \ldots, z_k$,*

$$S \vdash \bigwedge_{i \leq k} (z_i = x_i) \to \Phi\left(\ulcorner \bigwedge_{i \leq k} (\dot{z}_i = \dot{x}_i) \urcorner\right).$$

Proof. Since $T \vdash \bigwedge_{i \leq k} (z_i = z_i)$, we have

$$S \vdash \Phi\left(\ulcorner \bigwedge_{i \leq k} (\dot{z}_i = \dot{z}_i) \urcorner\right) \tag{2}$$

by $\mathbf{D1^U}$. Let $v_0, \ldots, v_k$ be fresh variables. By equality axioms of predicate calculus, we have

$$\mathsf{PA} \vdash \bigwedge_{i \leq k} (z_i = x_i) \to \left(\Phi\left(\ulcorner \bigwedge_{i \leq k} (\dot{v}_i = \dot{z}_i) \urcorner\right) \to \Phi\left(\ulcorner \bigwedge_{i \leq k} (\dot{v}_i = \dot{x}_i) \urcorner\right)\right).$$

By substituting z_i for v_i, we obtain

$$\mathsf{PA} \vdash \bigwedge_{i \leq k} (z_i = x_i) \to \left(\Phi\left(\ulcorner \bigwedge_{i \leq k} (\dot{z}_i = \dot{z}_i) \urcorner\right) \to \Phi\left(\ulcorner \bigwedge_{i \leq k} (\dot{z}_i = \dot{x}_i) \urcorner\right)\right).$$

By combining this with (2), we now obtain

$$S \vdash \bigwedge_{i \leq k} (z_i = x_i) \to \Phi\left(\ulcorner \bigwedge_{i \leq k} (\dot{z}_i = \dot{x}_i) \urcorner\right).$$

$\square$

For each term $t(\vec{x})$, let $c(t(\vec{x}))$ be the number of constant and function symbols contained in $t(\vec{x})$. We call $c(t(\vec{x}))$ the *complexity* of $t(\vec{x})$.

Lemma 3.5. *For any finite sequence $\{t_i(\vec{x})\}_{i \leq k}$ of terms with $\max_{i \leq k}\{c(t_i(\vec{x}))\} \leq 1$,*

$$S \vdash \bigwedge_{i \leq k} (z_i = t_i(\vec{x})) \to \Phi\left(\ulcorner \bigwedge_{i \leq k} (\dot{z}_i = t_i(\dot{\vec{x}})) \urcorner\right).$$

Proof. We prove by induction on the number m of terms of complexity 1 in such sequences. If a sequence does not contain terms of complexity 1, then it consists of variables, and hence the lemma holds for the sequence by Lemma 3.4.

Suppose that the lemma holds for such sequences with exactly m terms of complexity 1. Let $\{t_i(\vec{x})\}_{i\leq k}$ be any finite sequence consists of terms of complexity less than or equal to 1 and having exactly $m+1$ terms of complexity 1. We may assume that $c(t_k)=1$. Let $\xi(\vec{v}) :\equiv \bigwedge_{i<k}(z_i = t_i(\vec{x}))$. We distinguish the following four cases.

Case 1: $t_k(\vec{x})$ is 0. Then by induction hypothesis,

$$S \vdash \xi(\vec{v}) \wedge z_k = y \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \wedge \dot{z}_k = \dot{y} \urcorner).$$

By substituting 0 for y, we obtain

$$S \vdash \xi(\vec{v}) \wedge z_k = 0 \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \wedge \dot{z}_k = \dot{y} \urcorner)[0/y].$$

Since 0 is a numeral, we have

$$S \vdash \xi(\vec{v}) \wedge z_k = 0 \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \wedge \dot{z}_k = 0 \urcorner).$$

Case 2: $t_k(\vec{x})$ is $\mathsf{s}(x)$. By induction hypothesis,

$$S \vdash \xi(\vec{v}) \wedge z_k = y \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \wedge \dot{z}_k = \dot{y} \urcorner).$$

By substituting $\mathsf{s}(x)$ for y, we obtain

$$S \vdash \xi(\vec{v}) \wedge z_k = \mathsf{s}(x) \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \wedge \dot{z}_k = \dot{y} \urcorner)[\mathsf{s}(x)/y].$$

By Lemma 3.2, we conclude

$$S \vdash \xi(\vec{v}) \wedge z_k = \mathsf{s}(x) \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \wedge \dot{z}_k = \mathsf{s}(\dot{x}) \urcorner).$$

Case 3: $t_k(\vec{x})$ is $x+y$. Let $\varphi(y)$ be the formula

$$\forall x(\xi(\vec{v}) \wedge z_k = x + y \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \wedge \dot{z}_k = \dot{x} + \dot{y} \urcorner)).$$

By induction hypothesis,

$$S \vdash \xi(\vec{v}) \wedge z_k = x \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \wedge \dot{z}_k = \dot{x} \urcorner).$$

Since $\mathsf{PA} \vdash x = x+0$, we have $\mathsf{PA} \vdash (\xi(\vec{v}) \wedge z_k = x) \leftrightarrow (\xi(\vec{v}) \wedge z_k = x+0)$. Then by Lemma 3.1,

$$S \vdash \xi(\vec{v}) \wedge z_k = x + 0 \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \wedge \dot{z}_k = \dot{x} + 0 \urcorner).$$

This means $S \vdash \varphi(0)$.

By Lemma 3.2, we get

$$\mathsf{PA} \vdash \varphi(y) \wedge \xi(\vec{v}) \wedge z_k = \mathsf{s}(x) + y \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \wedge \dot{z}_k = \mathsf{s}(\dot{x}) + \dot{y} \urcorner).$$

Since $\mathsf{PA} \vdash \mathsf{s}(x) + y = x + \mathsf{s}(y)$, we obtain

$$S \vdash \varphi(y) \wedge \xi(\vec{v}) \wedge z_k = x + \mathsf{s}(y) \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \wedge \dot{z}_k = \dot{x} + \mathsf{s}(\dot{y}) \urcorner).$$

by Lemma 3.1. Then $S \vdash \varphi(y) \to \varphi(\mathsf{s}(y))$. By induction axiom, we conclude $S \vdash \forall y \varphi(y)$.

Case 4: $t_k(\vec{x})$ is $x \times y$. Let $\psi(y)$ be the formula

$$\forall w(\xi(\vec{v}) \wedge z_k = x \times y + w \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \wedge \dot{z}_k = \dot{x} \times \dot{y} + \dot{w} \urcorner)).$$

By induction hypothesis,

$$S \vdash \xi(\vec{v}) \wedge z_k = w \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \wedge \dot{z}_k = \dot{w} \urcorner).$$

Since $\mathsf{PA} \vdash w = x \times 0 + w$, we have

$$S \vdash \xi(\vec{v}) \wedge z_k = x \times 0 + w \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \wedge \dot{z}_k = \dot{x} \times 0 + \dot{w} \urcorner)$$

by Lemma 3.1. Therefore $S \vdash \psi(0)$.

Let $\rho(w)$ be the formula

$$\forall u(\xi(\vec{v}) \wedge z_k = x \times y + (u + w) \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \wedge \dot{z}_k = \dot{x} \times \dot{y} + (\dot{u} + \dot{w}) \urcorner)).$$

Then as in Case 3, we can prove $S \vdash \psi(y) \to \rho(0)$ and $S \vdash \rho(w) \to \rho(\mathsf{s}(w))$. Hence $S \vdash \psi(y) \to \forall w \rho(w)$. Then

$$S \vdash \psi(y) \wedge \xi(\vec{v}) \wedge z_k = x \times y + (x + w) \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \wedge \dot{z}_k = \dot{x} \times \dot{y} + (\dot{x} + \dot{w}) \urcorner).$$

Since $\mathsf{PA} \vdash x \times y + (x + w) = x \times \mathsf{s}(y) + w$, we get

$$S \vdash \psi(y) \wedge \xi(\vec{v}) \wedge z_k = x \times \mathsf{s}(y) + w \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \wedge \dot{z}_k = \dot{x} \times \mathsf{s}(\dot{y}) + \dot{w} \urcorner)$$

by Lemma 3.1. Thus $S \vdash \psi(y) \to \psi(\mathsf{s}(y))$, and hence $S \vdash \forall y \psi(y)$. By substituting 0 for w in $\psi(y)$, we obtain

$$S \vdash \xi(\vec{v}) \wedge z_k = x \times y + 0 \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \wedge \dot{z}_k = \dot{x} \times \dot{y} + 0 \urcorner).$$

Then the required conclusion follows from Lemma 3.1. $\square$

Lemma 3.6. *For any finite sequence $\{t_i(\vec{x})\}_{i \leq k}$ of terms,*

$$S \vdash \bigwedge_{i \leq k} (z_i = t_i(\vec{x})) \to \Phi\left(\ulcorner \bigwedge_{i \leq k} (\dot{z}_i = t_i(\dot{\vec{x}})) \urcorner\right).$$

Proof. We prove by induction on $\max_{i \leq k}\{c(t_i(\vec{x}))\}$. If $\max_{i \leq k}\{c(t_i(\vec{x}))\} \leq 1$, then the lemma follows from Lemma 3.5.

Suppose that the lemma holds for every finite sequence $\{t_i(\vec{x})\}_{i \leq k}$ of terms with $\max_{i \leq k}\{c(t_i(\vec{x}))\} = n \geq 1$. Then we show that the lemma holds for all finite sequences $\{t_i(\vec{x})\}_{i \leq k}$ containing only terms of complexity less than or equal to $n + 1$.

As in our proof of Lemma 3.5, this is proved by induction on the number m of terms of complexity $n + 1$ in such sequences. If $m = 0$, then the lemma follows from induction hypothesis. Then assume that the lemma holds for such sequences with exactly m terms of complexity $n+1$.

Let $\{t_i\}_{i \leq k}$ be any finite sequence consists of terms of complexity less than or equal to $n+1$ and having exactly $m+1$ terms of complexity $n+1$. We may assume that $c(t_k) = n+1$. Let $\xi(\vec{v}) :\equiv \bigwedge_{i<k}(z_i = t_i(\vec{x}))$. We give only a proof of the case that $t_k(\vec{x})$ is $\mathsf{s}(t'(\vec{x}))$ for some term $t'(\vec{x})$ of complexity n. Other cases are proved in a similar way.

Notice that $c(\mathsf{s}(w)) = 1 \leq n$ and $c(t'(\vec{x})) = n$. Then by induction hypothesis,

$$S \vdash \xi(\vec{v}) \land z_k = \mathsf{s}(w) \land w = t'(\vec{x}) \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \land \dot{z}_k = \mathsf{s}(\dot{w}) \land \dot{w} = t'(\dot{\vec{x}}) \urcorner).$$

Since $\mathsf{PA} \vdash \exists w(\xi(\vec{v}) \land z_k = \mathsf{s}(w) \land w = t'(\vec{x})) \leftrightarrow (\xi(\vec{v}) \land z_k = \mathsf{s}(t'(\vec{x})))$, we obtain

$$S \vdash \xi(\vec{v}) \land x_k = \mathsf{s}(t'(\vec{x})) \to \Phi(\ulcorner \xi(\dot{\vec{v}}) \land \dot{z}_k = \mathsf{s}(t'(\dot{\vec{x}})) \urcorner)$$

by Lemmas 3.4 and 3.1. □

Notice that each atomic formula $t_0 = t_1$ is equivalent to $\exists z(z = t_0 \land z = t_1)$, and each negated atomic formula $t_0 \neq t_1$ is PA-equivalent to $\exists z_0 \exists z_1(t_0 + \mathsf{s}(z_0) = t_1 \lor t_1 + \mathsf{s}(z_1) = t_0)$. Then we obtain the following lemma.

Lemma 3.7. *For any quantifier-free formula $\xi(\vec{x})$, there exists a quantifier-free formula $\delta(\vec{x}, \vec{y})$ satisfying the following conditions:*

1. $\mathsf{PA} \vdash \forall \vec{x}(\xi(\vec{x}) \leftrightarrow \exists \vec{y} \delta(\vec{x}, \vec{y}))$.
2. *$\delta(\vec{x}, \vec{y})$ is of the form $\delta_0(\vec{x}, \vec{y}) \lor \cdots \lor \delta_k(\vec{x}, \vec{y})$ and each disjunct $\delta_i(\vec{x}, \vec{y})$ is of the form*

$$\bigwedge_{j \leq l_i} (z_{i,j} = t_{i,j}(\vec{x}, \vec{y}))$$

for some terms $t_{i,0}(\vec{x}, \vec{y}), \ldots, t_{i,l_i}(\vec{x}, \vec{y})$ and variables $z_{i,0}, \ldots, z_{i,l_i} \in \vec{x}, \vec{y}$.

Also in our proof of Theorem 2.20, we use the following PA-provable form of the MRDP theorem.

Theorem 3.8 (The MRDP theorem (see [14])). *For any Σ_1 formula $\varphi(\vec{x})$, there exists a quantifier-free formula $\delta(\vec{x}, \vec{y})$ such that $\mathsf{PA} \vdash \forall \vec{x}(\varphi(\vec{x}) \leftrightarrow \exists \vec{y} \delta(\vec{x}, \vec{y}))$.*

Proof of Theorem 2.20. Let $\sigma(\vec{x})$ be any Σ_1 formula. We would like to prove $S \vdash \forall \vec{x}(\sigma(\vec{x}) \to \Phi(\ulcorner \sigma(\dot{\vec{x}}) \urcorner))$. By the MRDP theorem (Theorem 3.8), there exists a quantifier-free formula $\delta(\vec{x}, \vec{y})$ such that $\mathsf{PA} \vdash \forall \vec{x}(\sigma(\vec{x}) \leftrightarrow \exists \vec{y} \delta(\vec{x}, \vec{y}))$. By Lemma 3.7, we may assume that $\delta(\vec{x}, \vec{y})$ is of the form indicated in the statement of Lemma 3.7. For each $i \leq k$, by Lemma 3.6, we obtain

$$S \vdash \bigwedge_{j \leq l_i} (z_{i,j} = t_{i,j}(\vec{x}, \vec{y})) \to \Phi\left(\ulcorner \bigwedge_{j \leq l_i} (\dot{z}_{i,j} = t_{i,j}(\dot{\vec{x}}, \dot{\vec{y}})) \urcorner\right).$$

This means

$$S \vdash \delta_i(\vec{x}, \vec{y}) \to \Phi(\ulcorner \delta_i(\dot{\vec{x}}, \dot{\vec{y}}) \urcorner). \tag{3}$$

Since $\mathsf{PA} \vdash \delta_i(\vec{x}, \vec{y}) \to \delta(\vec{x}, \vec{y})$, $S \vdash \Phi(\ulcorner \delta_i(\dot{\vec{x}}, \dot{\vec{y}}) \urcorner) \to \Phi(\ulcorner \delta(\dot{\vec{x}}, \dot{\vec{y}}) \urcorner)$ by $\mathbf{B_2^U}$. Therefore by (3), $S \vdash \delta_i(\vec{x}, \vec{y}) \to \Phi(\ulcorner \delta(\dot{\vec{x}}, \dot{\vec{y}}) \urcorner)$. Since $i \leq k$ is arbitrary, we have $S \vdash \delta_0(\vec{x}, \vec{y}) \vee \cdots \vee \delta_k(\vec{x}, \vec{y}) \to \Phi(\ulcorner \delta(\dot{\vec{x}}, \dot{\vec{y}}) \urcorner)$. It follows $S \vdash \delta(\vec{x}, \vec{y}) \to \Phi(\ulcorner \delta(\dot{\vec{x}}, \dot{\vec{y}}) \urcorner)$. By Lemmas 3.4 and 3.1, we conclude $S \vdash \sigma(\vec{x}) \to \Phi(\ulcorner \sigma(\dot{\vec{x}}) \urcorner)$. $\square$

4 Witnesses for non-implications

In this section, we exhibit examples of formulas $\Phi(x)$ satisfying and not satisfying certain conditions. From these examples, several non-implications between conditions are concluded.

Our first two propositions give examples of formulas which do not satisfy **D1**. Proofs are easy and we omit them.

Proposition 4.1. *Let $\mathrm{Pr}_{\mathsf{Q}}(x)$ be the provability predicate of Robinson's arithmetic Q.*

1. $\mathrm{Pr}_{\mathsf{Q}}(x)$ *satisfies* $\mathbf{D2^G}$, $\mathbf{\Sigma_1C^G}$, $\mathbf{CB}$ *and* $\mathbf{PC^G}$.
2. $\mathrm{Pr}_{\mathsf{Q}}(x)$ *satisfies neither* $\mathbf{D1}$ *nor* $\mathbf{B_2}$.
3. $\mathsf{PA} \vdash \mathsf{Con}^H_{\mathrm{Pr}_{\mathsf{Q}}}$.

Proposition 4.2. *Let $\Psi(x) :\equiv x \neq x$.*

1. $\Psi(x)$ *satisfies* $\mathbf{D2^G}$, $\mathbf{D3^G}$, $\mathbf{B_2^U}$ *and* $\mathbf{CB}$.
2. $\Psi(x)$ *does not satisfy any of* $\mathbf{D1}$, $\mathbf{\Delta_0C}$ *and* $\mathbf{PC}$.
3. $\mathsf{PA} \vdash \mathsf{Con}^H_{\Psi}$.

Feferman [7] proved there exists a Π_1 numeration $\pi(v)$ of T in T such that $\mathsf{Con}^H_{\mathrm{Pr}_\pi}$ is provable in PA.

Fact 4.3 (Feferman [7])**.** *Suppose $S = T$.*

1. $\mathrm{Pr}_\pi(x)$ *is a Σ_2 provability predicate satisfying* $\mathbf{D1^U}$, $\mathbf{D2^G}$, $\mathbf{B_2^U}$, $\mathbf{\Sigma_1C^G}$, $\mathbf{CB}$ *and* $\mathbf{PC^G}$.
2. $\mathrm{Pr}_\pi(x)$ *does not satisfy* $\mathbf{D3}$.
3. $\mathsf{PA} \vdash \mathsf{Con}^H_{\mathrm{Pr}_\pi}$.

Mostowski (p. 24 in [20]) introduced the formula $\mathrm{Pr}^M_T(x) :\equiv \exists y(\mathrm{Prf}_T(x, y) \wedge \neg \mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, y))$ as an example of a Σ_1 provability predicate for which the second incompleteness theorem does not hold. Notice that $\mathrm{Pr}^M_T(x)$ is PA-provably equivalent to $\mathrm{Pr}_T(x) \wedge x \neq \ulcorner 0 \neq 0 \urcorner$ because $\mathsf{PA} \vdash \forall x_0 \forall x_1 \forall y(\mathrm{Prf}_T(x_0, y) \wedge \mathrm{Prf}_T(x_1, y) \to x_0 = x_1)$. The following proposition shows the situation for $\mathrm{Pr}^M_T(x)$.

Proposition 4.4.

1. $\mathrm{Pr}_T^M(x)$ *is a* Σ_1 *provability predicate satisfying* $\mathbf{D1^U}$, $\mathbf{\Sigma_1 C^G}$ *and* $\mathbf{PC^G}$.

2. $\mathrm{Pr}_T^M(x)$ *does not satisfy any of* $\mathbf{D2}$, $\mathbf{B_2}$ *and* $\mathbf{CB}$.

3. $\mathsf{PA} \vdash \mathsf{Con}^L_{\mathrm{Pr}_T^M}$ *and* $T \nvdash \mathsf{Con}^H_{\mathrm{Pr}_T^M}$.

The existence of Rosser provability predicates satisfying some derivability conditions were discussed by Bernardi and Montagna [4] and Arai [1]. They proved that there exists a Rosser provability predicate satisfying $\mathbf{D2^G}$. Also Arai proved the existence of a Rosser provability predicate satisfying $\mathbf{D3^G}$. Strictly speaking, in Arai's arguments, formulas are assumed to be in negation normal form (see [1]). We fix a natural algorithm calculating a negation normal form $\mathsf{nnf}(\varphi)$ of each formula φ satisfying $\mathsf{nnf}(\neg\neg\varphi) \equiv \mathsf{nnf}(\varphi)$. Then we can understand that Arai's Rosser provability predicates $\mathrm{Pr}^A(x)$ are of the form $\exists y(\mathrm{Prf}(\mathsf{nnf}(x), y) \land \forall z \leq y \neg \mathrm{Prf}(\mathsf{nnf}(\dot\neg x), z))$ for some suitable proof predicate $\mathrm{Prf}(x, y)$. Then $\mathsf{PA} \vdash \mathsf{Con}^H_{\mathrm{Pr}^A}$ always holds. Summarizing this observation, Arai's results are stated as follows.

Fact 4.5 (Arai [1])**.** *There exist* Σ_1 *provability predicates* $\mathrm{Pr}_1^A(x)$ *and* $\mathrm{Pr}_2^A(x)$ *of* T *with:*

1. $\mathrm{Pr}_1^A(x)$ *satisfies* $\mathbf{D1}$, $\mathbf{D2^G}$ *and* $\mathsf{PA} \vdash \mathsf{Con}^H_{\mathrm{Pr}_1^A}$.

2. $\mathrm{Pr}_2^A(x)$ *satisfies* $\mathbf{D1}$, $\mathbf{D3^G}$ *and* $\mathsf{PA} \vdash \mathsf{Con}^H_{\mathrm{Pr}_2^A}$.

By Proposition 2.4.4, $\mathrm{Pr}_1^A(x)$ satisfies $\mathbf{B_2}$. By Theorems 2.7 and 2.20, and Propositions 2.4, 2.13 and 2.14, $\mathrm{Pr}_1^A(x)$ does not satisfy any of $\mathbf{D1^U}$, $\mathbf{CB}$, $\mathbf{B_2^U}$, $\mathbf{D3}$ and $\mathbf{PC}$. By Theorems 2.8, 2.9 and 2.10 and Proposition 2.4.4, $\mathrm{Pr}_2^A(x)$ does not satisfy any of $\mathbf{D2}$, $\mathbf{B_2}$, $\mathbf{\Sigma_1 C}$ and $\mathbf{PC}$.

In [16], the author proved the existence of usual Rosser provability predicates satisfying additional derivability conditions. That is to say,

Fact 4.6 (Kurahashi [16])**.** *Suppose* $S = T$*. There exist* Σ_1 *provability predicates* $\mathrm{Pr}_1^R(x)$, $\mathrm{Pr}_2^R(x)$ *and* $\mathrm{Pr}_3^R(x)$ *of* T *with:*

1. $\mathrm{Pr}_1^R(x)$ *satisfies* $\mathbf{D1}$, $\mathbf{D2^G}$, $\mathbf{\Delta_0 C^G}$ *and* $\mathsf{PA} \vdash \mathsf{Con}^H_{\mathrm{Pr}_1^R}$.

2. $\mathrm{Pr}_2^R(x)$ *satisfies* $\mathbf{D1^U}$, $\mathbf{CB}$, $\mathbf{D2}$, $\mathbf{\Delta_0 C^G}$ *and* $\mathsf{PA} \vdash \mathsf{Con}^L_{\mathrm{Pr}_2^R}$.

3. $\mathrm{Pr}_3^R(x)$ *satisfies* $\mathbf{D1^U}$, $\mathbf{CB}$, $\mathbf{B_2}$, $\mathbf{D3^G}$, $\mathbf{\Delta_0 C^G}$ *and* $\mathsf{PA} \vdash \mathsf{Con}^L_{\mathrm{Pr}_3^R}$*, but does not satisfy* $\mathbf{\Sigma_1 C}$.

As in Fact 4.5.1, $\mathrm{Pr}_1^R(x)$ satisfies $\mathbf{B_2}$, but does not satisfy any of $\mathbf{D1^U}$, $\mathbf{CB}$, $\mathbf{B_2^U}$, $\mathbf{D3}$ and $\mathbf{PC}$. By Proposition 2.4.4, $\mathrm{Pr}_2^R(x)$ satisfies $\mathbf{B_2}$, but does not satisfy any of $\mathbf{D2^U}$, $\mathbf{D3}$, $\mathbf{B_2^U}$ and $\mathbf{PC}$ by Theorems 2.7 and 2.20, and Propositions 2.4.6 and 2.13.3. By Theorems 2.7 and 2.20 and Proposition 2.4, $\mathrm{Pr}_3^R(x)$ does not satisfy any of $\mathbf{D2}$, $\mathbf{B_2^U}$ and $\mathbf{PC}$.

In the remainder of this section, we introduce seven Σ_1 provability predicates $\mathrm{Pr}^{\mathrm{I}}_T(x)$, $\mathrm{Pr}^{\mathrm{II}}_T(x)$, $\mathrm{Pr}^{\mathrm{III}}_T(x)$, $\mathrm{Pr}^{\mathrm{IV}}_T(x)$, $\mathrm{Pr}^{\mathrm{V}}_T(x)$, $\mathrm{Pr}^{\mathrm{VI}}_T(x)$ and $\mathrm{Pr}^*(x)$ which indicate several non-implications of the conditions. The first three provability predicates are constructed in a similar way. Before introducing them, we prepare a definition and a lemma.

Definition 4.7. Let $\delta(x, z)$ be a Δ_1 formula.

1. $\mathrm{Prf}_T[\delta](x, y) :\equiv \mathrm{Prf}_T(x, y) \land \forall z < y(\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \to \delta(x, z))$.
2. $\mathrm{Pr}_T[\delta](x) :\equiv \exists y \mathrm{Prf}_T[\delta](x, y)$.

Lemma 4.8. *For any Δ_1 formula $\delta(x, z)$,*

1. *$\mathrm{Pr}_T[\delta](x)$ is a Σ_1 provability predicate of T.*
2. $\mathsf{PA} \vdash \forall x(\forall z(\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \to \delta(x, z)) \to (\mathrm{Pr}_T(x) \leftrightarrow \mathrm{Pr}_T[\delta](x)))$.
3. *If* $\mathsf{PA} \vdash \forall x \forall z(\mathsf{Fml}(x) \land x \leq z \to \delta(x, z))$, *then*

$$\mathsf{PA} \vdash \forall x \forall z(\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \land \mathsf{Fml}(x) \land \mathrm{Pr}_T[\delta](x) \to \delta(x, z)).$$

Proof. 1. Let φ be any formula and let n be any natural number. Since $\mathsf{PA} \vdash \forall z < \overline{n} \neg \mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z)$, $\mathsf{PA} \vdash \mathrm{Prf}_T(\ulcorner \varphi \urcorner, \overline{n}) \leftrightarrow \mathrm{Prf}_T[\delta](\ulcorner \varphi \urcorner, \overline{n})$. Since this equivalence is true in the standard model of arithmetic, we obtain that $\mathsf{PA} \vdash \mathrm{Pr}_T(\ulcorner \varphi \urcorner)$ if and only if $\mathsf{PA} \vdash \mathrm{Pr}_T[\delta](\ulcorner \varphi \urcorner)$. It follows that $\mathrm{Pr}_T[\delta](x)$ is also a Σ_1 provability predicate of T.

2. This is immediate from the definition.

3. Suppose $\mathsf{PA} \vdash \forall x \forall z(\mathsf{Fml}(x) \land x \leq z \to \delta(x, z))$. By the definition of $\mathrm{Prf}_T[\delta](x, y)$,

$$\mathsf{PA} \vdash \forall x \forall y \forall z(\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \land \mathrm{Prf}_T[\delta](x, y) \land z < y \to \delta(x, z)). \tag{4}$$

Since $\mathsf{PA} \vdash \mathrm{Prf}_T[\delta](x, y) \to \mathrm{Prf}_T(x, y)$ and $\mathsf{PA} \vdash \mathrm{Prf}_T(x, y) \to x \leq y$, we have $\mathsf{PA} \vdash \mathrm{Prf}_T[\delta](x, y) \to x \leq y$. Thus $\mathsf{PA} \vdash \mathrm{Prf}_T[\delta](x, y) \land y \leq z \to x \leq z$. By the supposition, $\mathsf{PA} \vdash \mathsf{Fml}(x) \land \mathrm{Prf}_T[\delta](x, y) \land y \leq z \to \delta(x, z)$. From this with (4), we obtain

$$\mathsf{PA} \vdash \forall x \forall y \forall z(\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \land \mathsf{Fml}(x) \land \mathrm{Prf}_T[\delta](x, y) \to \delta(x, z)),$$

and hence

$$\mathsf{PA} \vdash \forall x \forall z(\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \land \mathsf{Fml}(x) \land \mathrm{Pr}_T[\delta](x) \to \delta(x, z)).$$

$\square$

Let $\mathsf{Even}(x)$ be a natural Δ_1 formula saying that "x is the Gödel number of a formula containing an even number of logical symbols". Proposition 4.9 shows that full local derivability conditions do not imply uniform derivability conditions.

Proposition 4.9. *There exists a Σ_1 provability predicate $\mathrm{Pr}^{\mathrm{I}}_T(x)$ of T with:*

1. *$\mathrm{Pr}^{\mathrm{I}}_T(x)$ satisfies* **D1**, **D2** *and* $\mathbf{\Sigma_1 C}$.
2. *$\mathrm{Pr}^{\mathrm{I}}_T(x)$ does not satisfy any of* $\mathbf{D1^U}$, $\mathbf{D2^U}$, $\mathbf{D3^U}$, $\mathbf{\Delta_0 C^U}$ *and* $\mathbf{PC^U}$.

Proof. Let $\mathrm{Pr}^{\mathrm{I}}_T(x) :\equiv \mathrm{Pr}_T[x \leq z \lor \mathsf{Even}(x)](x)$. Then $\mathrm{Pr}^{\mathrm{I}}_T(x)$ is a Σ_1 provability predicate of T by Lemma 4.8.1. If $\mathrm{Pr}^{\mathrm{I}}_T(x)$ contains an even number of logical symbols, we replace $\mathrm{Pr}^{\mathrm{I}}_T(x)$ with $\mathrm{Pr}^{\mathrm{I}}_T(x) \land 0 = 0$. Then $\mathrm{Pr}^{\mathrm{I}}_T(x)$ contains an odd number of logical symbols, and hence $\mathsf{PA} \vdash \forall x \neg \mathsf{Even}(\ulcorner \mathrm{Pr}^{\mathrm{I}}_T(\dot{x}) \urcorner)$.

Let φ be any formula. Since $\mathsf{PA} \vdash \forall z(\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \to \ulcorner \varphi \urcorner \leq z \lor \mathsf{Even}(\ulcorner \varphi \urcorner))$, we have $\mathsf{PA} \vdash \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \leftrightarrow \mathrm{Pr}^{\mathrm{I}}_T(\ulcorner \varphi \urcorner)$ by Lemma 4.8.2. Therefore local derivability conditions for $\mathrm{Pr}^{\mathrm{I}}_T(x)$ are inherited from those for $\mathrm{Pr}_T(x)$.

We prove that $\mathrm{Pr}^{\mathrm{I}}_T(x)$ does not satisfy any of uniform derivability conditions. Since $\mathsf{PA} \vdash \forall x \forall z(\mathsf{Fml}(x) \land x \leq z \to (x \leq z \lor \mathsf{Even}(x)))$,

$$\mathsf{PA} \vdash \forall x \forall z(\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \land \mathsf{Fml}(x) \land \mathrm{Pr}^{\mathrm{I}}_T(x) \to (x \leq z \lor \mathsf{Even}(x)))$$

by Lemma 4.8.3. For the sake of simplicity, we deal with formulas whose only free variable is x. Let $\varphi(x)$ be such a formula. Then

$$\mathsf{PA} \vdash \forall x \forall z(\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \land \mathrm{Pr}^{\mathrm{I}}_T(\ulcorner \varphi(\dot{x}) \urcorner) \to (\ulcorner \varphi(\dot{x}) \urcorner \leq z \lor \mathsf{Even}(\ulcorner \varphi(\dot{x}) \urcorner))).$$

Since $\mathsf{PA} \vdash x \leq \ulcorner \varphi(\dot{x}) \urcorner$, we obtain

$$\mathsf{PA} \vdash \forall x \forall z(\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \land \mathrm{Pr}^{\mathrm{I}}_T(\ulcorner \varphi(\dot{x}) \urcorner) \to (x \leq z \lor \mathsf{Even}(\ulcorner \varphi(\dot{x}) \urcorner))). \tag{5}$$

- Since $\mathsf{PA} \vdash \forall x \neg \mathsf{Even}(\ulcorner 0 = 0 \land \dot{x} = \dot{x} \urcorner)$,

  $$\mathsf{PA} \vdash \forall x \forall z(\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \to (x \leq z \lor \neg \mathrm{Pr}^{\mathrm{I}}_T(\ulcorner 0 = 0 \land \dot{x} = \dot{x} \urcorner)))$$

  by (5). Hence $\mathsf{PA} \vdash \mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner) \to \exists x \neg \mathrm{Pr}^{\mathrm{I}}_T(\ulcorner 0 = 0 \land \dot{x} = \dot{x} \urcorner)$ because $\mathsf{PA} \vdash \forall z \exists x(x > z)$. It follows $S \nvdash \forall x \mathrm{Pr}^{\mathrm{I}}_T(\ulcorner 0 = 0 \land \dot{x} = \dot{x} \urcorner)$ because $S \nvdash \neg \mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner)$. This shows that $\mathrm{Pr}^{\mathrm{I}}_T(x)$ does not satisfy $\mathbf{D1^U}$.

- Let $\varphi(x)$ and $\psi(x)$ be formulas with $\mathsf{PA} \vdash \forall x \mathsf{Even}(\ulcorner \varphi(\dot{x}) \urcorner) \land \forall x \neg \mathsf{Even}(\ulcorner \psi(\dot{x}) \urcorner)$. Then $\mathsf{PA} \vdash \forall x \mathsf{Even}(\ulcorner \varphi(\dot{x}) \to \psi(\dot{x}) \urcorner)$. Since $\mathsf{PA} \vdash \mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner) \to \mathrm{Pr}_T(\ulcorner \varphi(\dot{x}) \to \psi(\dot{x}) \urcorner) \land \mathrm{Pr}_T(\ulcorner \varphi(\dot{x}) \urcorner)$, we have

  $$\mathsf{PA} \vdash \mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner) \to \mathrm{Pr}^{\mathrm{I}}_T(\ulcorner \varphi(\dot{x}) \to \psi(\dot{x}) \urcorner) \land \mathrm{Pr}^{\mathrm{I}}_T(\ulcorner \varphi(\dot{x}) \urcorner)$$

  by the choice of $\varphi(x)$ and $\psi(x)$, and the definition of $\mathrm{Prf}^{\mathrm{I}}_T(x, y)$. Suppose, towards a contradiction, that $\mathrm{Pr}^{\mathrm{I}}_T(x)$ satisfies $\mathbf{D2^U}$, then $S \vdash \mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner) \to \mathrm{Pr}^{\mathrm{I}}_T(\ulcorner \psi(\dot{x}) \urcorner)$. By (5), $S \vdash \mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \to (x \leq z \lor \mathsf{Even}(\ulcorner \psi(\dot{x}) \urcorner))$, and hence $S \vdash \mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner) \to \exists x \mathsf{Even}(\ulcorner \psi(\dot{x}) \urcorner)$. By the choice of $\psi(x)$, we obtain $S \vdash \neg \mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner)$. This is a contradiction. Therefore $\mathbf{D2^U}$ does not hold for $\mathrm{Pr}^{\mathrm{I}}_T(x)$.

- Let $\varphi(x)$ be a formula with $\mathsf{PA} \vdash \forall x \mathsf{Even}(\ulcorner \varphi(\dot{x}) \urcorner)$. Then $\mathsf{PA} \vdash \mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner) \to \mathrm{Pr}^{\mathrm{I}}_T(\ulcorner \varphi(\dot{x}) \urcorner)$ as described above. Suppose that $\mathbf{D3^U}$ holds for $\mathrm{Pr}^{\mathrm{I}}_T(x)$. Then $S \vdash \mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner) \to \mathrm{Pr}^{\mathrm{I}}_T(\ulcorner \mathrm{Pr}^{\mathrm{I}}_T(\ulcorner \varphi(\dot{x}) \urcorner) \urcorner)$. By (5), we have $S \vdash \mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner) \to \exists x \mathsf{Even}(\ulcorner \mathrm{Pr}^{\mathrm{I}}_T(\ulcorner \varphi(\dot{x}) \urcorner) \urcorner)$. Since $\mathrm{Pr}^{\mathrm{I}}_T(x)$ contains an odd number of logical symbols, $\neg \mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner)$ is proved in S, and this is a contradiction. Hence $\mathbf{D3^U}$ does not hold for $\mathrm{Pr}^{\mathrm{I}}_T(x)$.
- As described above, $\mathsf{PA} \vdash \mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner) \to \exists x \neg \mathrm{Pr}^{\mathrm{I}}_T(\ulcorner 0 = 0 \land \dot{x} = \dot{x} \urcorner)$. If $S \vdash \forall x(0 = 0 \land x = x \to \mathrm{Pr}^{\mathrm{I}}_T(\ulcorner 0 = 0 \land \dot{x} = \dot{x} \urcorner))$, then $S \vdash \mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner) \to \exists x \neg(0 = 0 \land x = x)$. This implies $S \vdash \neg \mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner)$, a contradiction. Therefore $S \nvdash \forall x(0 = 0 \land x = x \to \mathrm{Pr}^{\mathrm{I}}_T(\ulcorner 0 = 0 \land \dot{x} = \dot{x} \urcorner))$. This shows that $\mathbf{\Delta_0 C^U}$ does not hold for $\mathrm{Pr}^{\mathrm{I}}_T(x)$.
- $\mathbf{PC^U}$ fails to hold because $\mathsf{PA} \vdash \forall x \mathrm{Pr}_\emptyset(\ulcorner 0 = 0 \land \dot{x} = \dot{x} \urcorner)$.

□

By Proposition 2.4, $\mathrm{Pr}^{\mathrm{I}}_T(x)$ satisfies $\mathbf{B_2}$, $\mathbf{D3}$ and $\mathbf{PC}$. Propositions 2.13.1 and 2.14.1 imply that $\mathrm{Pr}^{\mathrm{I}}_T(x)$ satisfies neither $\mathbf{B_2^U}$ nor $\mathbf{CB}$.

Next we prove that full uniform derivability conditions do not imply any of global derivability conditions except for $\mathbf{D3^G}$, and that full derivability conditions are not sufficient for the unprovability of $\mathsf{Con}^{\Sigma_1}_{\Phi}$ even if $\Phi \in \Sigma_1$.

Proposition 4.10. *There exists a Σ_1 provability predicate $\mathrm{Pr}^{\mathrm{II}}_T(x)$ of T with:*

1. *$\mathrm{Pr}^{\mathrm{II}}_T(x)$ satisfies $\mathbf{D1^U}$, $\mathbf{D2^U}$, and $\mathbf{\Sigma_1 C^U}$.*
2. *$\mathrm{Pr}^{\mathrm{II}}_T(x)$ does not satisfy any of $\mathbf{D2^G}$, $\mathbf{\Delta_0 C^G}$ and $\mathbf{PC^G}$.*
3. *$\mathsf{PA} \vdash \mathsf{Con}^{\Sigma_1}_{\mathrm{Pr}^{\mathrm{II}}_T}$.*

Proof. For each formula φ, let $n(\varphi)$ be the number of occurrences of the symbol $\neg$ in φ. We may use a function symbol $n(x)$ corresponding to this function such that $\mathsf{PA} \vdash \forall x(\mathsf{Fml}(x) \to n(x) \leq x)$.

Let $\mathrm{Pr}^{\mathrm{II}}_T(x)$ be the Σ_1 formula $\mathrm{Pr}_T[n(x) \leq z \lor \mathsf{Even}(x)](x)$. Then $\mathrm{Pr}^{\mathrm{II}}_T(x)$ is a Σ_1 provability predicate of T by Lemma 4.8.1. Let $\varphi(\vec{x})$ be any formula. Then $\mathsf{PA} \vdash \forall \vec{x}(n(\ulcorner \varphi(\vec{\dot{x}}) \urcorner) = \overline{k})$ for some natural number k. Since $\mathsf{PA} \vdash \forall z(\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \to n(\ulcorner \varphi(\vec{\dot{x}}) \urcorner) \leq z \lor \mathsf{Even}(\ulcorner \varphi(\vec{\dot{x}}) \urcorner))$, we obtain $\mathsf{PA} \vdash \forall \vec{x}(\mathrm{Pr}_T(\ulcorner \varphi(\vec{\dot{x}}) \urcorner) \leftrightarrow \mathrm{Pr}^{\mathrm{II}}_T(\ulcorner \varphi(\vec{\dot{x}}) \urcorner))$ by Lemma 4.8.2. Therefore $\mathrm{Pr}^{\mathrm{II}}_T(x)$ satisfies $\mathbf{D1^U}$, $\mathbf{D2^U}$ and $\mathbf{\Sigma_1 C^U}$.

By Lemma 4.8.3, we have

$$\mathsf{PA} \vdash \forall x \forall z(\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \land \mathsf{Fml}(x) \land \mathrm{Pr}^{\mathrm{II}}_T(x) \to (n(x) \leq z \lor \mathsf{Even}(x))) \tag{6}$$

because $\mathsf{PA} \vdash \forall x(\mathsf{Fml}(x) \land x \leq z \to n(x) \leq z \lor \mathsf{Even}(x))$.

As in Proposition 4.9, failure of $\mathbf{D2^G}$, $\mathbf{\Delta_0 C^G}$ and $\mathbf{PC^G}$ for $\mathrm{Pr}^{\mathrm{II}}_T(x)$ follow from (6) and the facts $\mathsf{PA} \vdash \forall z \exists y(\mathsf{Fml}(y) \land n(y) > z \land \neg \mathsf{Even}(y))$, $\mathsf{PA} \vdash$

$\forall z \exists y(\mathsf{True}_{\Delta_0}(y) \land n(y) > z \land \neg\mathsf{Even}(y))$ and $\mathsf{PA} \vdash \forall z \exists y(\mathrm{Pr}_\emptyset(y) \land n(y) > z \land \neg\mathsf{Even}(y))$, respectively.

We prove $\mathsf{PA} \vdash \mathsf{Con}^{\Sigma_1}_{\mathrm{Pr}^{\mathrm{II}}_T}$. By (6) and $\mathsf{PA} \vdash \forall z \exists x(\Sigma_1(x) \land \mathsf{Sent}(x) \land n(x) > z \land \neg\mathsf{Even}(x))$, we have

$$\mathsf{PA} \vdash \forall z(\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \to \exists x(\Sigma_1(x) \land \mathsf{Sent}(x) \land \neg \mathrm{Pr}^{\mathrm{II}}_T(x))).$$

It follows $\mathsf{PA} \vdash \mathrm{Pr}^{\mathrm{II}}_T(\ulcorner 0 \neq 0 \urcorner) \to \mathsf{Con}^{\Sigma_1}_{\mathrm{Pr}^{\mathrm{II}}_T}$. On the other hand, obviously $\mathsf{PA} \vdash \neg\mathrm{Pr}^{\mathrm{II}}_T(\ulcorner 0 \neq 0 \urcorner) \to \mathsf{Con}^{\Sigma_1}_{\mathrm{Pr}^{\mathrm{II}}_T}$. Therefore we conclude $\mathsf{PA} \vdash \mathsf{Con}^{\Sigma_1}_{\mathrm{Pr}^{\mathrm{II}}_T}$. $\square$

From Propositions 2.13 and 2.14, $\mathrm{Pr}^{\mathrm{II}}_T(x)$ satisfies $\mathbf{B_2^U}$, $\mathbf{CB}$ and $\mathbf{PC^U}$. By Theorem 2.7, $T \nvdash \mathsf{Con}^{L}_{\mathrm{Pr}^{\mathrm{II}}_T}$.

We prove that the conditions $\Phi \in \Sigma_1$, $\mathbf{D1^U}$, $\mathbf{D2^G}$ and $\mathbf{\Sigma_1 C^G}$ are not sufficient for the unprovability of Gödel's consistency statement Con^{G}_{Φ}.

Proposition 4.11. *There exists a Σ_1 provability predicate $\mathrm{Pr}^{\mathrm{III}}_T(x)$ of T with:*

1. *$\mathrm{Pr}^{\mathrm{III}}_T(x)$ satisfies $\mathbf{D1^U}$, $\mathbf{D2^G}$ and $\mathbf{\Sigma_1 C^G}$.*
2. $\mathsf{PA} \vdash \mathsf{Con}^{G}_{\mathrm{Pr}^{\mathrm{III}}_T}$.

Proof. Let $\mathrm{Pr}^{\mathrm{III}}_T(x)$ be the formula $\mathrm{Pr}_T[\Sigma_z(x)](x)$. Then by Lemma 4.8.1, $\mathrm{Pr}^{\mathrm{III}}_T(x)$ is a Σ_1 provability predicate of T. For any formula $\varphi(\vec{x})$, we have $\mathsf{PA} \vdash \forall z \forall \vec{x}(\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \to \Sigma_z(\ulcorner \varphi(\dot{\vec{x}}) \urcorner))$ because $\mathsf{PA} \vdash \forall z \geq \overline{k}\Sigma_z(\ulcorner \varphi(\dot{\vec{x}}) \urcorner)$ for some natural number k. Hence $\mathsf{PA} \vdash \mathrm{Pr}_T(\ulcorner \varphi(\dot{\vec{x}}) \urcorner) \leftrightarrow \mathrm{Pr}^{\mathrm{III}}_T(\ulcorner \varphi(\dot{\vec{x}}) \urcorner)$ by Lemma 4.8.2. Thus $\mathbf{D1^U}$ holds for $\mathrm{Pr}^{\mathrm{III}}_T(x)$.

Since $\mathsf{PA} \vdash \forall x \forall z(\mathsf{Fml}(x) \land x \leq z \to \Sigma_z(x))$, we have

$$\mathsf{PA} \vdash \forall x \forall z(\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \land \mathsf{Fml}(x) \land \mathrm{Pr}^{\mathrm{III}}_T(x) \to \Sigma_z(x)) \tag{7}$$

by Lemma 4.8.3. Then

$$\mathsf{PA} \vdash \mathsf{Fml}(x) \land \mathsf{Fml}(y) \land \mathrm{Pr}^{\mathrm{III}}_T(x \dot{\to} y) \to (\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \to \Sigma_z(x \dot{\to} y)).$$

Thus

$$\mathsf{PA} \vdash \mathsf{Fml}(x) \land \mathsf{Fml}(y) \land \mathrm{Pr}^{\mathrm{III}}_T(x \dot{\to} y) \to \forall z(\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \to \Sigma_z(y)).$$

By Lemma 4.8.2,

$$\mathsf{PA} \vdash \mathsf{Fml}(x) \land \mathsf{Fml}(y) \land \mathrm{Pr}^{\mathrm{III}}_T(x \dot{\to} y) \to (\mathrm{Pr}_T(y) \leftrightarrow \mathrm{Pr}^{\mathrm{III}}_T(y)). \tag{8}$$

Since $\mathsf{PA} \vdash \mathrm{Pr}^{\mathrm{III}}_T(x \dot{\to} y) \land \mathrm{Pr}^{\mathrm{III}}_T(x) \to \mathrm{Pr}_T(x \dot{\to} y) \land \mathrm{Pr}_T(x)$, we have

$$\mathsf{PA} \vdash \mathsf{Fml}(x) \land \mathsf{Fml}(y) \land \mathrm{Pr}^{\mathrm{III}}_T(x \dot{\to} y) \land \mathrm{Pr}^{\mathrm{III}}_T(x) \to \mathrm{Pr}_T(y)$$

by $\mathbf{D2^G}$ for $\mathrm{Pr}_T(x)$. From this with (8),

$$\mathsf{PA} \vdash \mathsf{Fml}(x) \land \mathsf{Fml}(y) \land \mathrm{Pr}^{\mathrm{III}}_T(x \dot{\to} y) \land \mathrm{Pr}^{\mathrm{III}}_T(x) \to \mathrm{Pr}^{\mathrm{III}}_T(y).$$

This means $\mathbf{D2^G}$ holds for $\mathrm{Pr}_T^{\mathrm{III}}(x)$.

Since $\mathsf{PA} \vdash \mathsf{True}_{\Sigma_1}(x) \to \Sigma_1(x)$, $\mathsf{PA} \vdash \mathsf{True}_{\Sigma_1}(x) \to (\mathrm{Prf}_T(\ulcorner 0 \neq 0 \urcorner, z) \to \Sigma_z(x))$. By Lemma 4.8.2, $\mathsf{PA} \vdash \mathsf{True}_{\Sigma_1}(x) \to (\mathrm{Pr}_T(x) \leftrightarrow \mathrm{Pr}_T^{\mathrm{III}}(x))$. By $\mathbf{\Sigma_1 C^G}$ for $\mathrm{Pr}_T(x)$, we obtain $\mathsf{PA} \vdash \mathsf{True}_{\Sigma_1}(x) \to \mathrm{Pr}_T^{\mathrm{III}}(x)$.

By (7) and $\mathsf{PA} \vdash \forall z \exists x(\mathsf{Fml}(x) \land \neg \Sigma_z(x))$, we have $\mathsf{PA} \vdash \mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner) \to \exists x(\mathsf{Fml}(x) \land \neg \mathrm{Pr}_T^{\mathrm{III}}(x))$. Thus $\mathsf{PA} \vdash \mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner) \to \mathsf{Con}^G_{\mathrm{Pr}_T^{\mathrm{III}}}$. On the other hand, since $\mathsf{PA} \vdash \neg \mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner) \to \neg \mathrm{Pr}_T^{\mathrm{III}}(\ulcorner 0 \neq 0 \urcorner)$, we have $\mathsf{PA} \vdash \neg \mathrm{Pr}_T(\ulcorner 0 \neq 0 \urcorner) \to \mathsf{Con}^G_{\mathrm{Pr}_T^{\mathrm{III}}}$. Therefore $\mathsf{PA} \vdash \mathsf{Con}^G_{\mathrm{Pr}_T^{\mathrm{III}}}$. $\square$

By Propositions 2.13 and 2.14, $\mathrm{Pr}_T^{\mathrm{III}}(x)$ satisfies $\mathbf{B_2^U}$, $\mathbf{CB}$ and $\mathbf{PC^U}$. Corollary 2.28 implies that $\mathbf{PC^G}$ fails to hold for $\mathrm{Pr}_T^{\mathrm{III}}(x)$ and $T \nvdash \mathsf{Con}^{\Sigma_1}_{\mathrm{Pr}_T^{\mathrm{III}}}$.

We prove that there exists a Σ_1 provability predicate which satisfies the Hilbert–Bernays–Löb derivability conditions, but does not satisfy $\mathbf{\Sigma_1 C}$. The following proof is based on the construction presented in Section 5 of Visser [24].

Proposition 4.12. *There exists a Σ_1 provability predicate $\mathrm{Pr}_T^{\mathrm{IV}}(x)$ of T which satisfies $\mathbf{D1}$, $\mathbf{D2^G}$ and $\mathbf{D3^G}$, but does not satisfy $\mathbf{\Sigma_1 C}$.*

Proof. We say an $\mathcal{L}_A$-formula φ is *propositionally atomic* if it is not a Boolean combination of proper subformulas of φ. We fix a bijective mapping f from the set of all propositional variables to the set of all propositionally atomic formulas. For each propositionally atomic formula $\Phi(x)$, the mapping f can be extended to the mapping f_Φ from the set of all modal formulas to the set of all $\mathcal{L}_A$-formulas satisfying the following clauses:

1. $f_\Phi(p)$ is $f(p)$ for each propositional variable p;

2. f_Φ commutes with every propositional connective;

3. $f_\Phi(\Box A)$ is $\Phi(\ulcorner f_\Phi(A) \urcorner)$.

For any finite set X of modal formulas and any modal formula A, A is said to be derived in X if A is provable in the system whose axioms are elements of X and whose inference rules are Modus Ponens $\dfrac{B \quad B \to C}{C}$ and Necessitation $\dfrac{B}{\Box B}$.

For each natural number n, let $\mathrm{Th}_n(T)$ be the finite set of all $\mathcal{L}_A$-formulas having a T-proof whose Gödel number is less than or equal to n. We write $T \vdash_{\Phi,n} \varphi$ if there exist a finite set X of modal formulas and a modal formula A such that $f_\Phi(X) = \mathrm{Th}_n(T)$, $f_\Phi(A)$ is φ and A is derived in X. For $m < n$, $T \vdash_{\Phi,m} \varphi$ implies $T \vdash_{\Phi,n} \varphi$ because $\mathrm{Th}_m(T) \subseteq \mathrm{Th}_n(T)$. As shown in Visser [24], the ternary relation $T \vdash_{\Phi,n} \varphi$ is computable. Thus we obtain a Δ_1 formula $P_T(\ulcorner \Phi \urcorner, x, y)$ saying that x is the Gödel number of a formula φ satisfying $T \vdash_{\Phi,y} \varphi$.

By the Fixed Point Lemma, there exist a Σ_1 formula $\mathrm{Pr}_T^{\mathrm{IV}}(x)$ and a Σ_1 sentence σ satisfying the following equivalences:

1. $P'_T(x,y) \equiv P_T(\ulcorner \mathrm{Pr}_T^{\mathrm{IV}} \urcorner, x, y)$;
2. $\mathsf{PA} \vdash \mathrm{Pr}_T^{\mathrm{IV}}(x) \leftrightarrow \exists y(P'_T(x,y) \land \forall z < y \neg P'_T(\ulcorner \neg\sigma \urcorner, z))$;
3. $\mathsf{PA} \vdash \sigma \leftrightarrow \exists z(P'_T(\ulcorner \neg\sigma \urcorner, z) \land \forall y \leq z \neg P'_T(\ulcorner \sigma \urcorner, y))$.

First, we prove $T \nvdash_{\mathrm{Pr}_T^{\mathrm{IV}}, n} \neg\sigma$ for all n by induction on n. Suppose $T \nvdash_{\mathrm{Pr}_T^{\mathrm{IV}}, m} \neg\sigma$ for all $m < n$. Then $\mathsf{PA} \vdash \forall z < \overline{n} \neg P'_T(\ulcorner \neg\sigma \urcorner, z)$.

Let X be any finite set of modal formulas with $f_{\mathrm{Pr}_T^{\mathrm{IV}}}(X) = \mathrm{Th}_n(T)$. Let A be any modal formula derived in X, then $T \vdash_{\mathrm{Pr}_T^{\mathrm{IV}}, n} f_{\mathrm{Pr}_T^{\mathrm{IV}}}(A)$. Hence we have $\mathsf{PA} \vdash P'_T(\ulcorner f_{\mathrm{Pr}_T^{\mathrm{IV}}}(A) \urcorner, \overline{n})$, and thus $\mathsf{PA} \vdash \mathrm{Pr}_T^{\mathrm{IV}}(\ulcorner f_{\mathrm{Pr}_T^{\mathrm{IV}}}(A) \urcorner)$. Moreover, we show $T \vdash f_{\mathrm{Pr}_T^{\mathrm{IV}}}(A)$. This is proved by induction on the length of derivation in X. If $A \in X$, then $f_{\mathrm{Pr}_T^{\mathrm{IV}}}(A) \in \mathrm{Th}_n(T)$, and $f_{\mathrm{Pr}_T^{\mathrm{IV}}}(A)$ has a T-proof. If A is derived from B and $B \to A$ by Modus Ponens and $T \vdash f_{\mathrm{Pr}_T^{\mathrm{IV}}}(B) \land f_{\mathrm{Pr}_T^{\mathrm{IV}}}(B \to A)$, then $T \vdash f_{\mathrm{Pr}_T^{\mathrm{IV}}}(A)$. If A is derived from B by Necessitation, then A is of the form $\Box B$. Since $\mathsf{PA} \vdash \mathrm{Pr}_T^{\mathrm{IV}}(\ulcorner f_{\mathrm{Pr}_T^{\mathrm{IV}}}(B) \urcorner)$ as above, we get $\mathsf{PA} \vdash f_{\mathrm{Pr}_T^{\mathrm{IV}}}(A)$. In this paragraph, we have shown that if $T \vdash_{\mathrm{Pr}_T^{\mathrm{IV}}, n} \varphi$, then $T \vdash \varphi$.

Suppose, towards a contradiction, $T \vdash_{\mathrm{Pr}_T^{\mathrm{IV}}, n} \neg\sigma$. Then $T \vdash \neg\sigma$. Since $T \nvdash \sigma$, $T \nvdash_{\mathrm{Pr}_T^{\mathrm{IV}}, m} \sigma$ for all $m \leq n$. Therefore $\mathsf{PA} \vdash P'_T(\ulcorner \neg\sigma \urcorner, \overline{n}) \land \forall y \leq \overline{n} \neg P'_T(\ulcorner \sigma \urcorner, y)$. By the definition of σ, we have $\mathsf{PA} \vdash \sigma$. This is a contradiction. We obtain $T \nvdash_{\mathrm{Pr}_T^{\mathrm{IV}}, n} \neg\sigma$.

If $T \vdash \varphi$, then $\varphi \in \mathrm{Th}_n(T)$ for some n. Then $T \vdash_{\mathrm{Pr}_T^{\mathrm{IV}}, n} \varphi$ trivially holds, and hence $\mathsf{PA} \vdash P'_T(\ulcorner \varphi \urcorner, \overline{n})$. Since $\mathsf{PA} \vdash \forall z < \overline{n} P'_T(\ulcorner \neg\sigma \urcorner, z)$, we obtain $\mathsf{PA} \vdash \mathrm{Pr}_T^{\mathrm{IV}}(\ulcorner \varphi \urcorner)$. On the other hand, we assume $\mathsf{PA} \vdash \mathrm{Pr}_T^{\mathrm{IV}}(\ulcorner \varphi \urcorner)$. Then $P'_T(\ulcorner \varphi \urcorner, \overline{n})$ is true in the standard model of arithmetic for some n. This means $T \vdash_{\mathrm{Pr}_T^{\mathrm{IV}}, n} \varphi$. Then we obtain $T \vdash \varphi$. Therefore we have shown that $\mathrm{Pr}_T^{\mathrm{IV}}(x)$ is a Σ_1 provability predicate of T.

We prove $\mathbf{D2^G}$ for $\mathrm{Pr}_T^{\mathrm{IV}}(x)$. We work in S. Suppose $\mathrm{Pr}_T^{\mathrm{IV}}(\ulcorner \varphi \urcorner)$ and $\mathrm{Pr}_T^{\mathrm{IV}}(\ulcorner \varphi \to \psi \urcorner)$ are true. Then for some n, $T \vdash_{\mathrm{Pr}_T^{\mathrm{IV}}, n} \varphi$, $T \vdash_{\mathrm{Pr}_T^{\mathrm{IV}}, n} \varphi \to \psi$ and $T \nvdash_{\mathrm{Pr}_T^{\mathrm{IV}}, m} \neg\sigma$ for all $m < n$. Then $T \vdash_{\mathrm{Pr}_T^{\mathrm{IV}}, n} \psi$. Thus $\mathrm{Pr}_T^{\mathrm{IV}}(\ulcorner \psi \urcorner)$ is true.

We prove $\mathbf{D3^G}$ for $\mathrm{Pr}_T^{\mathrm{IV}}(x)$. We proceed in S. Suppose $\mathrm{Pr}_T^{\mathrm{IV}}(\ulcorner \varphi \urcorner)$ is true. Then for some n, $T \vdash_{\mathrm{Pr}_T^{\mathrm{IV}}, n} \varphi$ and $T \nvdash_{\mathrm{Pr}_T^{\mathrm{IV}}, m} \neg\sigma$ for all $m < n$. Then $T \vdash_{\mathrm{Pr}_T^{\mathrm{IV}}, n} \mathrm{Pr}_T^{\mathrm{IV}}(\ulcorner \varphi \urcorner)$. Thus $\mathrm{Pr}_T^{\mathrm{IV}}(\ulcorner \mathrm{Pr}_T^{\mathrm{IV}}(\ulcorner \varphi \urcorner) \urcorner)$ is true.

At last, we prove that $\mathbf{\Sigma_1 C}$ fails to hold. Suppose, for a contradiction, $T \vdash \sigma \to \mathrm{Pr}_T^{\mathrm{IV}}(\ulcorner \sigma \urcorner)$. By witness comparison argument, we have $\mathsf{PA} \vdash \sigma \to \neg\mathrm{Pr}_T^{\mathrm{IV}}(\ulcorner \sigma \urcorner)$. Thus $T \vdash \neg\sigma$. Then $T \vdash_{\mathrm{Pr}_T^{\mathrm{IV}}, n} \neg\sigma$ for some n. This is a contradiction. Therefore we conclude $T \nvdash \sigma \to \mathrm{Pr}_T^{\mathrm{IV}}(\ulcorner \sigma \urcorner)$. $\Box$

By Proposition 2.4, Theorem 2.20, Proposition 2.13.3 and Proposition 2.14.1, $\mathrm{Pr}_T^{\mathrm{IV}}(x)$ does not satisfy any of $\mathbf{PC}$, $\mathbf{B_2^U}$, $\mathbf{D1^U}$ and $\mathbf{CB}$.

The next two propositions show that $\{\mathbf{D1}, \mathbf{\Sigma_1 C}\}$ and $\{\mathbf{D1}, \mathbf{PC}\}$ are incomparable.

Proposition 4.13. *There exists a Σ_1 provability predicate $\mathrm{Pr}_T^{\mathrm{V}}(x)$ of T which satisfies $\mathbf{\Sigma_1 C^G}$, but does not satisfy any of $\mathbf{D1^U}$ and $\mathbf{PC}$.*

Proof. Let T_0 be any finite subtheory of T containing Q with $\bigwedge T_0$ is not a Π_1 sentence. Let $\mathrm{Prf}'_T(v, x, y)$ be the Δ_1 formula

$$\mathrm{Prf}_T(x, y) \land (\exists z < y \mathrm{Prf}_T(\dot{\neg} v, z) \to \Sigma_1(x)).$$

By the Fixed Point Lemma, there exists a Σ_1 sentence σ satisfying

$$\mathsf{PA} \vdash \sigma \leftrightarrow \exists z(\mathrm{Prf}_T(\ulcorner \neg\sigma \urcorner, z) \land \forall y \leq z \neg \mathrm{Prf}'_T(\ulcorner \sigma \urcorner, \ulcorner \bigwedge T_0 \to \sigma \urcorner, y)).$$

Let $\mathrm{Prf}^{\mathrm{V}}_T(x, y) :\equiv \mathrm{Prf}'_T(\ulcorner \sigma \urcorner, x, y)$ and let $\mathrm{Pr}^{\mathrm{V}}_T(x) :\equiv \exists y \mathrm{Prf}^{\mathrm{V}}_T(x, y)$. Then

- $\mathsf{PA} \vdash \mathrm{Prf}^{\mathrm{V}}_T(x, y) \leftrightarrow \mathrm{Prf}_T(x, y) \land (\exists z < y \mathrm{Prf}_T(\ulcorner \neg\sigma \urcorner, z) \to \Sigma_1(x))$.
- $\mathsf{PA} \vdash \sigma \leftrightarrow \exists z(\mathrm{Prf}_T(\ulcorner \neg\sigma \urcorner, z) \land \forall y \leq z \neg \mathrm{Prf}^{\mathrm{V}}_T(\ulcorner \bigwedge T_0 \to \sigma \urcorner, y))$.

First, we prove $T \nvdash \neg\sigma$. If $T \vdash \neg\sigma$, then for some natural number p, $\mathsf{PA} \vdash \mathrm{Prf}_T(\ulcorner \neg\sigma \urcorner, \overline{p})$. Since $T \nvdash \sigma$, obviously $T \nvdash \bigwedge T_0 \to \sigma$. Then $\mathsf{PA} \vdash \forall y \leq \overline{p} \neg \mathrm{Prf}_T(\ulcorner \bigwedge T_0 \to \sigma \urcorner, y)$. Since $\mathrm{Prf}^{\mathrm{V}}_T(x, y)$ implies $\mathrm{Prf}_T(x, y)$, we have $\mathsf{PA} \vdash \forall y \leq \overline{p} \neg \mathrm{Prf}^{\mathrm{V}}_T(\ulcorner \bigwedge T_0 \to \sigma \urcorner, y)$. Then $\mathsf{PA} \vdash \sigma$ by the definition of σ. This is a contradiction. Therefore $T \nvdash \neg\sigma$.

It follows that for any natural number n, $\mathsf{PA} \vdash \neg \mathrm{Prf}_T(\ulcorner \neg\sigma \urcorner, \overline{n})$. Then for any formula φ, $\mathsf{PA} \vdash \mathrm{Prf}_T(\ulcorner \varphi \urcorner, \overline{n}) \leftrightarrow \mathrm{Prf}^{\mathrm{V}}_T(\ulcorner \varphi \urcorner, \overline{n})$. Thus $\mathrm{Pr}^{\mathrm{V}}_T(x)$ is a Σ_1 provability predicate of T.

Since $\mathsf{PA} \vdash \Sigma_1(x) \to (\mathrm{Pr}_T(x) \leftrightarrow \mathrm{Pr}^{\mathrm{V}}_T(x))$ by the definition, $\mathbf{\Sigma_1 C^G}$ for $\mathrm{Pr}^{\mathrm{V}}_T(x)$ easily follows from $\mathbf{\Sigma_1 C^G}$ for $\mathrm{Pr}_T(x)$.

We prove that $\mathbf{PC}$ fails to hold for $\mathrm{Pr}^{\mathrm{V}}_T(x)$. If $\mathrm{Pr}^{\mathrm{V}}_T(x)$ satisfied $\mathbf{PC}$, then $S \vdash \mathrm{Pr}_\emptyset(\ulcorner \bigwedge T_0 \to \sigma \urcorner) \to \mathrm{Pr}^{\mathrm{V}}_T(\ulcorner \bigwedge T_0 \to \sigma \urcorner)$. By formalized deduction theorem, $S \vdash \mathrm{Pr}_{[T_0]}(\ulcorner \sigma \urcorner) \to \mathrm{Pr}^{\mathrm{V}}_T(\ulcorner \bigwedge T_0 \to \sigma \urcorner)$. By $\mathbf{\Sigma_1 C}$ for $\mathrm{Pr}_{[T_0]}(x)$,

$$S \vdash \sigma \to \mathrm{Pr}^{\mathrm{V}}_T(\ulcorner \bigwedge T_0 \to \sigma \urcorner). \tag{9}$$

By the definition of $\mathrm{Prf}^{\mathrm{V}}_T(x, y)$, we obtain

$$\mathsf{PA} \vdash \mathrm{Prf}^{\mathrm{V}}_T(\ulcorner \bigwedge T_0 \to \sigma \urcorner, y) \land \mathrm{Prf}_T(\ulcorner \neg\sigma \urcorner, z) \land z < y \to \Sigma_1(\ulcorner \bigwedge T_0 \to \sigma \urcorner).$$

Since $\bigwedge T_0 \to \sigma$ is not Σ_1,

$$\mathsf{PA} \vdash \mathrm{Prf}^{\mathrm{V}}_T(\ulcorner \bigwedge T_0 \to \sigma \urcorner, y) \land \mathrm{Prf}_T(\ulcorner \neg\sigma \urcorner, z) \to y \leq z.$$

It follows

$$\mathsf{PA} \vdash \mathrm{Pr}^{\mathrm{V}}_T(\ulcorner \bigwedge T_0 \to \sigma \urcorner) \to \forall z(\mathrm{Prf}_T(\ulcorner \neg\sigma \urcorner, z) \to \exists y \leq z \mathrm{Prf}^{\mathrm{V}}_T(\ulcorner \bigwedge T_0 \to \sigma \urcorner, y)).$$

This means $\mathsf{PA} \vdash \mathrm{Pr}^{\mathrm{V}}_T(\ulcorner \bigwedge T_0 \to \sigma \urcorner) \to \neg\sigma$. From this with (9), $S \vdash \sigma \to \neg\sigma$, and hence $S \vdash \neg\sigma$. This is a contradiction. Therefore $\mathrm{Pr}^{\mathrm{V}}_T(x)$ does not satisfy $\mathbf{PC}$.

Finally, we prove that $\mathrm{Pr}^{\mathrm{V}}_T(x)$ does not satisfy $\mathbf{D1^U}$. Let $\varphi(x)$ be any formula such that $\mathsf{PA} \vdash \forall x \neg \Sigma_1(\ulcorner \varphi(\dot{x}) \urcorner)$ and $T \vdash \forall x \varphi(x)$. Since $\mathsf{PA} \vdash \mathrm{Prf}_T(\ulcorner \varphi(\dot{z}) \urcorner, y) \to$

$z < y$, we have $\mathsf{PA} \vdash \mathrm{Pr}_T^{\mathrm{V}}(\ulcorner \varphi(\dot{z}) \urcorner) \land \mathrm{Prf}_T(\ulcorner \neg\sigma \urcorner, z) \to \Sigma_1(\ulcorner \varphi(\dot{z}) \urcorner)$ by the definition of $\mathrm{Prf}_T^{\mathrm{V}}(x, y)$. Hence $\mathsf{PA} \vdash \mathrm{Pr}_T^{\mathrm{V}}(\ulcorner \varphi(\dot{z}) \urcorner) \to \neg\mathrm{Prf}_T(\ulcorner \neg\sigma \urcorner, z)$. Then $\mathsf{PA} \vdash \forall x \mathrm{Pr}_T^{\mathrm{V}}(\ulcorner \varphi(\dot{x}) \urcorner) \to \neg\mathrm{Pr}_T(\ulcorner \neg\sigma \urcorner)$. Since $T \nvdash \neg\mathrm{Pr}_T(\ulcorner \neg\sigma \urcorner)$, we conclude that $T \nvdash \forall x \mathrm{Pr}_T^{\mathrm{V}}(\ulcorner \varphi(\dot{x}) \urcorner)$. □

By Propositions 2.4 and 2.14. $\mathrm{Pr}_T^{\mathrm{V}}(x)$ does not satisfy any of $\mathbf{D2}$, $\mathbf{B_2}$ and $\mathbf{CB}$.

We give an example of Mostowski-like Σ_1 provability predicate which satisfies $\mathbf{PC^G}$ but does not satisfy $\mathbf{\Sigma_1 C}$.

Proposition 4.14. *There exists a Σ_1 provability predicate $\mathrm{Pr}_T^{\mathrm{VI}}(x)$ of T with:*

1. *$\mathrm{Pr}_T^{\mathrm{VI}}(x)$ satisfies $\mathbf{D1^U}$, $\mathbf{D3^G}$, $\mathbf{\Delta_0 C^G}$ and $\mathbf{PC^G}$.*
2. *$\mathrm{Pr}_T^{\mathrm{VI}}(x)$ satisfies neither $\mathbf{\Sigma_1 C}$ nor $\mathbf{CB}$.*

Proof. Let ξ be a Π_1 sentence undecidable in T such as Rosser's sentence (see [17]), and let ξ' be the sentence $\xi \lor 0 = \mathsf{s}(0)$ which is also undecidable in T. Let $\mathrm{Pr}_T^{\mathrm{VI}}(x) :\equiv \mathrm{Pr}_T(x) \land x \neq \ulcorner \neg\xi' \urcorner$. Obviously,

$$\mathsf{PA} \vdash \forall x (x \neq \ulcorner \neg\xi' \urcorner \to (\mathrm{Pr}_T(x) \leftrightarrow \mathrm{Pr}_T^{\mathrm{VI}}(x))). \tag{10}$$

Since $\neg\xi'$ is not provable in T, $\mathrm{Pr}_T^{\mathrm{VI}}(x)$ is a Σ_1 provability predicate of T, and also $\mathbf{D1^U}$ holds for $\mathrm{Pr}_T^{\mathrm{VI}}(x)$. The conditions $\mathbf{D3^G}$ and $\mathbf{\Delta_0 C^G}$ follow from $\mathsf{PA} \vdash \forall x (\ulcorner \mathrm{Pr}_T^{\mathrm{VI}}(\dot{x}) \urcorner \neq \ulcorner \neg\xi' \urcorner)$ and $\mathsf{PA} \vdash \forall x (\mathsf{True}_{\Delta_0}(x) \to x \neq \ulcorner \neg\xi' \urcorner)$, respectively.

We prove $\mathbf{PC^G}$. Let M be an $\mathcal{L}_A$-structure whose domain is a singleton $\{e\}$. Then for every closed $\mathcal{L}_A$-term t, $t^M = e$. Thus $M \models \xi \lor 0 = \mathsf{s}(0)$. Therefore $\neg\xi'$ is not provable in predicate calculus. The above argument can be formalized in PA, and so $\mathsf{PA} \vdash \forall x (\mathsf{Fml}(x) \to (\mathrm{Pr}_\emptyset(x) \to x \neq \ulcorner \neg\xi' \urcorner))$. Then by $\mathbf{PC^G}$ for $\mathrm{Pr}_T(x)$, we conclude $\mathsf{PA} \vdash \forall x (\mathsf{Fml}(x) \to (\mathrm{Pr}_\emptyset(x) \to \mathrm{Pr}_T^{\mathrm{VI}}(x)))$.

Since $\mathsf{PA} \vdash \neg\mathrm{Pr}_T^{\mathrm{VI}}(\ulcorner \neg\xi' \urcorner)$ and $T \nvdash \xi'$, we can prove $S \nvdash \mathrm{Pr}_T^{\mathrm{VI}}(\ulcorner \forall x \neg(\xi \lor x = \mathsf{s}(0)) \urcorner) \to \forall x \mathrm{Pr}_T^{\mathrm{VI}}(\ulcorner \neg(\xi \lor \dot{x} = \mathsf{s}(0)) \urcorner)$ by (10). The conditions $\mathbf{\Sigma_1 C}$ and $\mathbf{CB}$ fail to hold because of them. □

By Proposition 2.4, $\mathrm{Pr}_T^{\mathrm{VI}}(x)$ satisfies neither $\mathbf{D2}$ nor $\mathbf{B_2}$.

At last, we prove that our Theorem 2.20 is actually an improvement of Buchholz's theorem (Theorem 2.18).

Theorem 4.15. *There exists a Σ_1 provability predicate $\mathrm{Pr}^*(x)$ of PA which satisfies $\mathbf{D1^U}$, $\mathbf{B_2^U}$, $\mathbf{\Sigma_1 C^G}$ and $\mathbf{PC^G}$ but does not satisfy $\mathbf{D2}$.*

This theorem is proved by using Beklemishev's arithmetical completeness theorem of the bimodal logic CS_2 with respect to independent Σ_1 numerations (see Beklemishev [3]). For this, we need some preparations. The language of CS_2 is that of propositional logic equipped with two unary modal operators [0] and [1]. Formulas in this language are called CS_2-formulas. The axioms of the bimodal logic CS_2 are propositional tautologies and the formulas $[i](p \to q) \to ([i]p \to [i]q)$, $[i]p \to [j][i]p$ and $[i]([i]p \to p) \to [i]p$ for $i, j \in \{0, 1\}$. The inference rules of CS_2 are modus ponens $\dfrac{A, \quad A \to B}{B}$, necessitation $\dfrac{A}{[i]A}$ for $i \in \{0,1\}$, and uniform substitution.

We say a structure $M = (W, K_0, K_1, \prec, \Vdash, b)$ is a CS_2-*model* if it satisfies the following conditions:

1. W is a nonempty finite set.
2. K_0 and K_1 are subsets of W with $W = K_0 \cup K_1$.
3. $\prec$ is a strict partial ordering over W.
4. $b \in K_0 \cap K_1$ and $b \prec x$ for all $x \in W \setminus \{b\}$.
5. $\Vdash$ is a binary relation between W and the set of all CS_2-formulas such that $\Vdash$ satisfies the usual conditions for satisfaction and the following condition: for $i \in \{0,1\}$, $x \Vdash [i]A$ if and only if for all $y \in K_i$, if $x \prec y$, then $y \Vdash A$.

A CS_2-formula A is said to be *true* in a CS_2-model $M = (W, K_0, K_1, \prec, \Vdash, b)$ if $b \Vdash A$. The modal logic CS_2 is sound and complete with respect to CS_2 models.

Theorem 4.16 (See Smoryński [22]). *For any CS_2-formula A, the following are equivalent:*

1. $\mathsf{CS}_2 \vdash A$.
2. *A is true in all CS_2-models.*

Let $\alpha_0(v)$ and $\alpha_1(v)$ be any Σ_1 numerations of PA. A mapping f from CS_2-formulas to $\mathcal{L}_A$-sentences is a (α_0, α_1)-*interpretation* if f commutes with each propositional connective, and $f([i]A) \equiv \mathrm{Pr}_{\alpha_i}(\ulcorner f(A) \urcorner)$ for $i \in \{0,1\}$. Beklemishev proved that CS_2 is sound and complete with respect to this kind of interpretations.

Theorem 4.17 (The arithmetical completeness theorem of CS_2 (Beklemishev [3])). *For any CS_2-formula A, the following are equivalent:*

1. $\mathsf{CS}_2 \vdash A$.
2. *For any Σ_1 numerations $\alpha_0(v)$ and $\alpha_1(v)$ of PA and any (α_0, α_1)-interpretation f, $\mathsf{PA} \vdash f(A)$.*

We are ready to prove Theorem 4.15.

Proof of Theorem 4.15. Let us consider a CS_2-model $M = (W, K_0, K_1, \prec, \Vdash, b)$ satisfying the following conditions:

1. $W = \{b, x_0, x_1\}$,
2. $K_0 = \{b, x_0\}$ and $K_1 = \{b, x_1\}$,
3. $\prec = \{(b, x_0), (b, x_1)\}$,

4. $x_0 \Vdash p$ and $x_1 \nVdash p$.

Then $b \Vdash [0]p \wedge [1]\neg p \wedge \neg[0]\bot \wedge \neg[1]\bot$. Thus $\mathsf{CS_2} \nvdash [0]p \wedge [1]\neg p \to [0]\bot \vee [1]\bot$. By the arithmetical completeness theorem of $\mathsf{CS_2}$, there are Σ_1 numerations $\alpha_0(v)$ and $\alpha_1(v)$ of PA, and a (α_0, α_1)-interpretation f such that $\mathsf{PA} \nvdash f([0]p \wedge [1]\neg p \to [0]\bot \vee [1]\bot)$. Let $\xi :\equiv f(p)$, then

$$\mathsf{PA} \nvdash \mathrm{Pr}_{\alpha_0}(\ulcorner \xi \urcorner) \wedge \mathrm{Pr}_{\alpha_1}(\ulcorner \neg\xi \urcorner) \to \neg\mathsf{Con}_{\mathrm{Pr}_{\alpha_0}} \vee \neg\mathsf{Con}_{\mathrm{Pr}_{\alpha_1}}. \tag{11}$$

Let $\mathrm{Pr}^*(x)$ be the Σ_1 formula $\mathrm{Pr}_{\alpha_0}(x) \vee \mathrm{Pr}_{\alpha_1}(x)$. Then $\mathrm{Pr}^*(x)$ is obviously a Σ_1 provability predicate of PA. Moreover $\mathbf{D1^U}$, $\mathbf{\Sigma_1 C^G}$ and $\mathbf{PC^G}$ are inherited from $\mathrm{Pr}_{\alpha_0}(x)$.

First, we prove that $\mathrm{Pr}^*(x)$ satisfies $\mathbf{B_2^U}$. Suppose $\mathsf{PA} \vdash \forall \vec{x}(\varphi(\vec{x}) \to \psi(\vec{x}))$. Then since both $\mathrm{Pr}_{\alpha_0}(x)$ and $\mathrm{Pr}_{\alpha_1}(x)$ satisfy $\mathbf{B_2^U}$, we have

$$\mathsf{PA} \vdash \mathrm{Pr}_{\alpha_0}(\ulcorner \varphi(\dot{\vec{x}}) \urcorner) \to \mathrm{Pr}_{\alpha_0}(\ulcorner \psi(\dot{\vec{x}}) \urcorner) \text{ and } \mathsf{PA} \vdash \mathrm{Pr}_{\alpha_1}(\ulcorner \varphi(\dot{\vec{x}}) \urcorner) \to \mathrm{Pr}_{\alpha_1}(\ulcorner \psi(\dot{\vec{x}}) \urcorner).$$

By the definition of $\mathrm{Pr}^*(x)$,

$$\mathsf{PA} \vdash \mathrm{Pr}_{\alpha_0}(\ulcorner \varphi(\dot{\vec{x}}) \urcorner) \to \mathrm{Pr}^*(\ulcorner \psi(\dot{\vec{x}}) \urcorner) \text{ and } \mathsf{PA} \vdash \mathrm{Pr}_{\alpha_1}(\ulcorner \varphi(\dot{\vec{x}}) \urcorner) \to \mathrm{Pr}^*(\ulcorner \psi(\dot{\vec{x}}) \urcorner).$$

Therefore we conclude

$$\mathsf{PA} \vdash \forall \vec{x}(\mathrm{Pr}^*(\ulcorner \varphi(\dot{\vec{x}}) \urcorner) \to \mathrm{Pr}^*(\ulcorner \psi(\dot{\vec{x}}) \urcorner)).$$

At last, we prove that $\mathrm{Pr}^*(x)$ does not satisfy $\mathbf{D2}$. Suppose, towards a contradiction,

$$\mathsf{PA} \vdash \mathrm{Pr}^*(\ulcorner \xi \to 0 \neq 0 \urcorner) \to (\mathrm{Pr}^*(\ulcorner \xi \urcorner) \to \mathrm{Pr}^*(\ulcorner 0 \neq 0 \urcorner)).$$

Then by the definition of $\mathrm{Pr}^*(x)$,

$$\mathsf{PA} \vdash \mathrm{Pr}_{\alpha_0}(\ulcorner \neg\xi \urcorner) \vee \mathrm{Pr}_{\alpha_1}(\ulcorner \neg\xi \urcorner) \to (\mathrm{Pr}_{\alpha_0}(\ulcorner \xi \urcorner) \vee \mathrm{Pr}_{\alpha_1}(\ulcorner \xi \urcorner) \to \neg\mathsf{Con}_{\mathrm{Pr}_{\alpha_0}} \vee \neg\mathsf{Con}_{\mathrm{Pr}_{\alpha_1}}).$$

By logic, we obtain

$$\mathsf{PA} \vdash \mathrm{Pr}_{\alpha_0}(\ulcorner \xi \urcorner) \wedge \mathrm{Pr}_{\alpha_1}(\ulcorner \neg\xi \urcorner) \to \neg\mathsf{Con}_{\mathrm{Pr}_{\alpha_0}} \vee \neg\mathsf{Con}_{\mathrm{Pr}_{\alpha_1}}.$$

This contradicts (11). Therefore we conclude

$$\mathsf{PA} \nvdash \mathrm{Pr}^*(\ulcorner \xi \to 0 \neq 0 \urcorner) \to (\mathrm{Pr}^*(\ulcorner \xi \urcorner) \to \mathrm{Pr}^*(\ulcorner 0 \neq 0 \urcorner)).$$

$\square$

By Proposition 2.14.2, $\mathrm{Pr}^*(x)$ satisfies $\mathbf{CB}$.

As we have seen, examples of formulas given in this section show several non-implications between conditions. For instance, the following non-implications related to Proposition 2.4 are also obtained.

1. $\mathbf{\Delta_0 C} \not\Rightarrow \mathbf{D1}$ (Proposition 4.1).

2. $\{\mathbf{B_m} : m \geq 2\} \not\Rightarrow \mathbf{D1}$ (Proposition 4.2).
   For all $m \geq 2$, $\mathbf{D1} \not\Rightarrow \mathbf{B_m}$ (Proposition 4.4).

3. For all $m \geq 1$, $\mathbf{D2} \not\Rightarrow \mathbf{B_m}$ (Proposition 4.1).

4. $\mathbf{D3} \not\Rightarrow \mathbf{\Delta_0 C}$ (Proposition 4.2).

However, we do not have enough such non-implications between conditions including uniform and global versions. We close this paper with the following problem.

Problem 4.18. *Study further non-implications between derivability conditions.*

5 Acknowledgments

This work was partly supported by JSPS KAKENHI Grant Numbers 16K17653 and 19K14586. The author would like to thank Toshiyasu Arai, Yong Cheng, Makoto Kikuchi, Hidenori Kurokawa, Yuya Okawa and Albert Visser for their valuable comments. The author would also like to thank the anonymous referee for pointing out an error in an earlier version of the manuscript.

References

[1] Toshiyasu Arai. Derivability conditions on Rosser's provability predicates. *Notre Dame Journal of Formal Logic*, 31(4):487–497, 1990.

[2] Sergei N. Artemov and Lev D. Beklemishev. *Provability Logic*, volume 13 of *Handbook of Philosophical Logic*, pages 189–360. Springer, Dordrecht, 2nd edition, 2005.

[3] Lev D. Beklemishev. Independent enumerations of theories and recursive progressions. *Sibirskii Matematichskii Zhurnal*, 33(5):22–46, 1992. English translation is in *Siberian Mathematical Journal*, 33(5), 760-783, 1992.

[4] Claudio Bernardi and Franco Montagna. Equivalence relations induced by extensional formulae: classification by means of a new fixed point property. *Fundamenta Mathematicae*, 124(3):221–233, 1984.

[5] George Boolos. *The logic of provability*. Cambridge University Press, Cambridge, 1993.

[6] Wilfried Buchholz. Mathematische Logik II. `http://www.mathematik.uni-muenchen.de/~buchholz/articles/LogikII.ps`, 1993.

[7] Solomon Feferman. Arithmetization of metamathematics in a general setting. *Fundamenta Mathematicae*, 49:35–92, 1960.

[8] Kurt Gödel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. (in German). *Monatshefte für Mathematik und Physik*, 38(1):173–198, 1931. English translation in Kurt Gödel, *Collected Works*, Vol. 1 (pp. 145–195).

[9] Petr Hájek and Pavel Pudlák. *Metamathematics of First-Order Arithmetic*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin, 1993.

[10] David Hilbert and Paul Bernays. *Grundlagen der Mathematik. Vol. II*. Springer, Berlin, 1939.

[11] G. E. Hughes and M. J. Cresswell. *A new introduction to modal logic*. Routledge, London, 1996.

[12] Giorgi Japaridze and Dick de Jongh. *The logic of provability*, volume 137 of *Studies in Logic and the Foundations of Mathematics*, pages 475–546. North-Holland, Amsterdam, 1998.

[13] Robert G. Jeroslow. Redundancies in the Hilbert-Bernays derivability conditions for Gödel's second incompleteness theorem. *The Journal of Symbolic Logic*, 38(3):359–367, 1973.

[14] Richard Kaye. *Models of Peano arithmetic*, volume 15 of *Oxford Logic Guides*. Oxford Science Publications, New York, 1991.

[15] Gerog Kreisel and Gaisi Takeuti. Formally self-referential propositions for cut free classical analysis and related systems. *Dissertationes Mathematicae (Rozprawy Matematyczne)*, 118, 1974.

[16] Taishi Kurahashi. Rosser provability and the second incompleteness theorem. In *Symposium on Advances in Mathematical Logic 2018 proceedings*, 2020. Accepted.

[17] Per Lindström. *Aspects of Incompleteness*. Number 10 in Lecture Notes in Logic. A K Peters, 2nd edition, 2003.

[18] Martin Hugo Löb. Solution of a problem of Leon Henkin. *The Journal of Symbolic Logic*, 20(2):115–118, 1955.

[19] Franco Montagna. On the formulas of Peano arithmetic which are provably closed under modus ponens. *Bollettino dell'Unione Matematica Italiana*, 16(B5):196–211, 1979.

[20] Andrzej Mostowski. Thirty years of foundational studies: lectures on the development of mathematical logic and the study of the foundations of mathematics in 1930-1964. In *Acta Philosophica Fennica*, volume 17, pages 1–180. 1965.

[21] Wolfgang Rautenberg. *A concise introduction to mathematical logic. Third edition*. Universitext. Springer, New York, 2010.

[22] Craig Smoryński. *Self-reference and modal logic*. Universitext. Springer-Verlag, New York, 1985.

[23] Alfred Tarski, Andrzej Mostowski, and Raphael M. Robinson. *Undecidable Theories*, volume 13 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing, 1953.

[24] Albert Visser. Transductions in arithmetic. *Annals of Pure and Applied Logic*, 167(3):211–234, 2016.

[25] Albert Visser. The absorption law, or how to Kreisel a Hilbert-Bernays-Löb. arXiv: 1804.07465, 2018.

[26] Christopher von Bülow. A remark on equivalent Rosser sentences. *Annals of Pure and Applied Logic*, 151(1):62–67, 2008.