Gödel’s Theorem Fails for $\Pi_1$ Axiomatizations

Michael Wehar
University at Buffalo
mwehar@buffalo.edu

December 28, 2013

Abstract

We introduce a $\Pi_1$ set $S$ for which Gödel’s Second Incompleteness Theorem fails. In particular, we show $\text{ZF} \vdash \text{Con}(\text{ZF} + \text{Con}(\text{ZF})) \rightarrow \text{Con}(S) \land \text{Pf}_S(\text{Con}(S))$. Then, we carefully analyze the relationship between $\text{Pf}_S(x)$ and $\text{Pf}_S(\text{Pf}_S(x))$ in order to show $\text{ZF} \vdash \text{Con}(\text{ZF} + \text{Con}(\text{ZF})) \rightarrow \exists x [\text{Pf}_S(x) \land \neg \text{Pf}_S(\text{Pf}_S(x))]$.

1 Preliminaries

Definition 1.1. Let $G$ denote the set of Gödel numbers for well-formed formulas of the first order system for Zermelo-Fraenkel Set Theory.

Definition 1.2. Let $\text{ZF} \subseteq G$ denote the set of Gödel numbers for the standard axioms of Zermelo-Fraenkel Set Theory.

Definition 1.3. For all $A \subseteq G$ and $\Gamma X \in G$, let $\text{Pf}_A(\Gamma X)$ express that there is a proof of the well-formed formula represented by $\Gamma X$ from the set of formulas represented by members of $A$. For the remainder of the paper, we will omit the corner brackets and write $\text{Pf}_A(X)$ to improve readability.

We will take for granted that $\text{Pf}$ can be recursively defined within the first order system for Zermelo-Fraenkel Set Theory. In addition, at the top level we write $\text{ZF} \vdash W$ to express that one could present a formal proof of the well-formed formula $W$ using the first order system for Zermelo-Fraenkel Set Theory.

Definition 1.4. Let $\text{Con}(A)$ abbreviate $\neg \text{Pf}_A(0 = 1)$.
We will leave the following propositions as exercises for the reader.

**Proposition 1.1.** $\text{ZF} \vdash \forall A, B [A \subseteq B \rightarrow \forall x [\text{Pf}_A(x) \rightarrow \text{Pf}_B(x)]]$.

**Proposition 1.2.** $\text{ZF} \vdash \forall A, B [A \subseteq B \land \text{Con}(B) \rightarrow \text{Con}(A)]$.

**Proposition 1.3.** $\text{ZF} \vdash \forall A \forall x [\text{Con}(A) \land \text{Pf}_A(x) \rightarrow \text{Con}(A + x)]$.

**Proposition 1.4.** $\text{ZF} \vdash \forall A [\exists x \text{Pf}_A(\neg x \land x) \leftrightarrow \forall x \text{Pf}_A(x)]$.

In addition, we will make use of the following well-known theorems.

**Deduction Theorem.** $\text{ZF} \vdash \forall A \forall x, y [\text{Pf}_{A+x}(y) \leftrightarrow \text{Pf}_A(x \rightarrow y)]$.

**Diagonal Lemma.** For every formula $p(x)$, there exists a sentence $\psi \in \mathcal{G}$ such that

$$\text{ZF} \vdash \psi \leftrightarrow p(\uparrow \psi \downarrow).$$

**Gödel’s Second Incompleteness Theorem.** Let a $\Sigma_1$ formula $\phi(x)$ be given. Let $T$ denote the set associated with $\phi(x)$. If $\text{ZF} \vdash \text{ZF} \subseteq T \subseteq \mathcal{G}$, then

$$\text{ZF} \vdash \text{Pf}_T(\text{Con}(T)) \rightarrow \neg \text{Con}(T).$$

The requirement on $T$ being defined by a $\Sigma_1$ formula $\phi(x)$ is significant. It is necessary that $\phi(x)$ is embedded in the proof. See Appendix 4.1 for more details.

## 2 Gödel’s Theorem Fails

Consider the following extension\(^1\) of ZF:

$$S := \begin{cases} 
\text{ZF} + \text{Con}(\text{ZF}) & \text{if } \text{Con}(\text{ZF} + \text{Con}(\text{ZF})) \\
\text{ZF} & \text{otherwise.}
\end{cases}$$

$S$ is $\Pi_1$ because there is a program that enumerates the complement. This follows because ZF is decidable and we can determine if $\text{Con}(\text{ZF}) \notin S$ by searching for a proof of $0 = 1$ with axioms from $\text{ZF} + \text{Con}(\text{ZF})$.

If one could prove that $\text{Con}(\text{ZF})$ implies $\text{Con}(\text{ZF} + \text{Con}(\text{ZF}))$, then ZF proves its own inconsistency. However, we will prove in the following that $\text{Con}(\text{ZF})$ implies $\text{Con}(S)$.

---

\(^1\)Formally, one could define $S$ using pairing, comprehension, and union.
Lemma 2.1. ZF ⊨ Con(S) ↔ Con(ZF).

Proof. The claim follows from the following three statements using the method of proof by cases.

a) ZF ⊨ S = ZF + Con(ZF) → [Con(S) ↔ Con(ZF)]

b) ZF ⊨ S = ZF → [Con(S) ↔ Con(ZF)]

c) ZF ⊨ S = ZF + Con(ZF) ∨ S = ZF.

First, we show a.

\[ S = ZF + \text{Con}(ZF) \implies \text{Con}(ZF + \text{Con}(ZF)) \]  \hspace{1cm} (1)
\[ \implies \text{Con}(S) \land \text{Con}(ZF) \] \hspace{1cm} (2)
\[ \implies \text{Con}(S) \leftrightarrow \text{Con}(ZF). \] \hspace{1cm} (3)

(1) follows from the definition of S.
(2) follows from proposition 1.2 because ZF ⊆ S ⊆ ZF + Con(ZF).
(3) follows from logical axioms.

Lastly, b follows from the axioms for equality and c follows from the definition of S and logical axioms. \qed

Theorem 2.1. ZF ⊨ Con(ZF + Con(ZF)) → Con(S) ∧ Pf_{S}(\text{Con}(S)).

Proof. First, by proposition 1.2, we have ZF ⊨ Con(ZF + Con(ZF)) → Con(S) because S ⊆ ZF + Con(ZF). Next, we show ZF ⊨ Con(ZF + Con(ZF)) → Pf_{S}(\text{Con}(S)).

\[ \text{Con}(ZF + \text{Con}(ZF)) \implies S = ZF + \text{Con}(ZF) \] \hspace{1cm} (4)
\[ \implies Pf_{S}(\text{Con}(ZF)) \] \hspace{1cm} (5)
\[ \implies Pf_{S}(\text{Con}(S)). \] \hspace{1cm} (6)

(4) follows from the definition of S.
(5) follows because Con(ZF) ∈ S.
(6) One can use the proof of lemma 2.1 to show ZF ⊨ Pf_{ZF}(\text{Con}(S) ↔ Con(ZF)). Since ZF ⊆ S, we can apply proposition 1.1 to get ZF ⊨ Pf_{S}(\text{Con}(S) ↔ Con(ZF)). \qed
3 Incompatible Proof Levels

We introduced a set $S$ whose members depend on a property that is potentially independent of ZF. In particular, $\text{Con}(ZF) \in S$ if and only if $\text{Con}(ZF + \text{Con}(ZF))$. We will show that if $\text{Con}(ZF) \in S$, then $S$ has incompatible proof levels, that is $S$ proves the sentence $\text{Con}(S)$, but does not prove that it proves $\text{Con}(S)$.

Lemma 3.1. $ZF \vdash \text{Con}(S) \land \text{Pf}_S(\text{Con}(S)) \rightarrow \text{Con}(ZF + \text{Con}(ZF))$.

Proof.

$$\text{Con}(S) \land \text{Pf}_S(\text{Con}(S)) \Rightarrow \text{Con}(S) \land \text{Pf}_S(\text{Con}(ZF)) \quad (7)$$
$$\Rightarrow \text{Con}(S + \text{Con}(ZF)) \quad (8)$$
$$\Rightarrow \text{Con}(ZF + \text{Con}(ZF)). \quad (9)$$

$$(7)$$ One can use the proof of lemma 2.1 to show $ZF \vdash \text{Pf}_{ZF}(\text{Con}(S) \leftrightarrow \text{Con}(ZF))$. Since $ZF \subseteq S$, we can apply proposition 1.1 to get $ZF \vdash \text{Pf}_S(\text{Con}(S) \leftrightarrow \text{Con}(ZF))$.

$$(8)$$ follows from proposition 1.3.

$$(9)$$ follows because $S + \text{Con}(ZF) = ZF + \text{Con}(ZF)$. $\square$

Theorem 3.1. $ZF \vdash \text{Con}(ZF + \text{Con}(ZF)) \rightarrow \neg \text{Pf}_S(\text{Pf}_S(\text{Con}(S)))$.

Proof.

$$\text{Con}(ZF + \text{Con}(ZF)) \Rightarrow \neg \text{Pf}_{ZF + \text{Con}(ZF)}(\text{Con}(ZF + \text{Con}(ZF))) \quad (10)$$
$$\Rightarrow \neg \text{Pf}_S(\text{Con}(ZF + \text{Con}(ZF))) \quad (11)$$
$$\Rightarrow \neg \text{Pf}_S(\text{Con}(S) \land \text{Pf}_S(\text{Con}(S))) \quad (12)$$
$$\Rightarrow \neg \text{Pf}_S(\text{Pf}_S(\text{Con}(S))). \quad (13)$$

$$(10)$$ follows from Gödel’s Second Incompleteness Theorem.

$$(11)$$ follows from proposition 1.1 because $S \subseteq ZF + \text{Con}(ZF)$.

$$(12)$$ One can use the proof of lemma 3.1 to show

$$ZF \vdash \text{Pf}_{ZF}(\text{Con}(S) \land \text{Pf}_S(\text{Con}(S)) \rightarrow \text{Con}(ZF + \text{Con}(ZF))).$$

Since $ZF \subseteq S$, we can apply proposition 1.1 to get

$$ZF \vdash \text{Pf}_S(\text{Con}(S) \land \text{Pf}_S(\text{Con}(S)) \rightarrow \text{Con}(ZF + \text{Con}(ZF))).$$
(13) follows from theorem 2.1 because $\text{ZF} \vdash \text{Con}(\text{ZF} + \text{Con}(\text{ZF})) \rightarrow \text{Pf}_S(\text{Con}(S))$.  

\begin{corollary}
\text{Corollary 3.1.} $\text{ZF} \vdash \text{Con}(\text{ZF} + \text{Con}(\text{ZF})) \rightarrow \exists x [\text{Pf}_S(x) \land \neg \text{Pf}_S(\text{Pf}_S(x))].$
\end{corollary}

\begin{corollary}
\text{Corollary 3.2.} $\text{ZF}$ proves that the following are equivalent:

\begin{enumerate}
\item $\neg \text{Con}(\text{ZF} + \text{Con}(\text{ZF}))$
\item $\text{Pf}_{\text{ZF}}(\neg \text{Con}(\text{ZF}))$
\item $\text{Pf}_S(\neg \text{Con}(S))$
\item $\text{Pf}_S(\text{Pf}_S(\text{Con}(S))).$
\end{enumerate}
\end{corollary}

\section{Appendix}

\subsection{Gödel’s Second Incompleteness Theorem}

Gödel’s Theorem is a theorem scheme. If a $\Sigma_1$ formula $\phi(x)$ is provided, then one could carry out the proof. The assumption that $\phi(x)$ is $\Sigma_1$ is significant. Since $\phi(x)$ is $\Sigma_1$, $T$ is computably enumerable and one could write a program $p$ that enumerates codings of $T$-proofs \textit{i.e.} proofs whose axioms are from $T$. Therefore, if $t$ is a coding of a $T$-proof, then there is a computation for $p$ that accepts $t$. The existence of a computation implies the existence of a proof that $t$ is in fact a $T$-proof. Since $t$ is an arbitrary $T$-proof, one could formalize the preceding to get

$$\text{ZF} \vdash \forall x [\text{Pf}_T(x) \rightarrow \text{Pf}_T(\text{Pf}_T(x))]$$

which is needed to carry out the proof that $T$ proves its own consistency implies $T$ is inconsistent.

There is a terrible subtlety in the preceding discussion. We require that a $\Sigma_1$ formula $\phi(x)$ is provided. In particular, we cannot generalize over all $\Sigma_1$ sets $T$. Pick two distinct $\Sigma_1$ sets $Y_1$ and $Y_2$. Consider the following set

$$W := \begin{cases} 
Y_1 & \text{if CH} \\
Y_2 & \text{otherwise.}
\end{cases}$$

If CH denotes the Continuum Hypothesis, then whether $W$ is associated with $Y_1$’s formula or $Y_2$’s formula is independent of ZF. Therefore, we cannot provide a $\Sigma_1$ formula
for W. Formally, Gödel’s Theorem will not apply to W because we need to use the $\Sigma_1$ formula to prove “that $t$ is in fact a $T$-proof”, as stated above.

4.2 Known Results for Complete Theories

It is worth noting that complete extensions of ZF are known to have properties similar to S. In particular, if one defines a complete extension $T$ of ZF, then we observe that $\text{ZFC} \vdash \text{Con}(T + \text{Con}(T)) \rightarrow \text{Con}(T) \land \text{Pf}_T(\text{Con}(T))$ and using the Diagonal Lemma can show $\text{ZFC} \vdash \text{Con}(T + \text{Con}(T)) \rightarrow \exists x [\text{Pf}_T(x) \land \neg \text{Pf}_T(\text{Pf}_T(x))]$.

5 Acknowledgments

I greatly appreciate all of the help and suggestions that I received. In particular, I would like to thank Joseph Swernofsky, James Clay, Kyle Russ, Travis Nell, Matthew Mirman, Kenneth Regan, Richard Lipton, and Leo Harrington.

References

