Gödel’s Incompleteness Theorems Interpreted through the Superposition Principle

Lee Chul (commple@gmail.com)

This paper was written by training ChatGPT on the Principle of Superposition of Opposition and Negation, which the author established through the study of ancient philosophy, and then using ChatGPT as a writing assistant.

Abstract

This paper interprets Gödel's incompleteness theorems through the Superposition Principle. The central question is not merely what Gödel's theorem states, but why a sufficiently strong formal system becomes incomplete when it attempts to internalize its own conditions of judgment. The paper begins from the first axiom of the Superposition Principle: an item and its opposite are not two independently given substances, but two opposite manifestations of a single oppositional pair within a field of judgment. Applied to truth and falsehood, this means that truth and falsehood remain distinct at the truth-value level, T != F, while they are structurally related within a truth-falsehood pair P_T = (T/F). The notation T ~ F therefore does not mean the collapse of classical truth values into T = F. It means that truth and falsehood are mutually determined as opposite manifestations of one judgmental structure. The paper then introduces the Principle of Superposition of Opposition and Negation. A binary opposition P_A = (A/A*) opens a fourfold structure Q_A = {A, A*, ¬_C A, ¬_C A*} when context-dependent negative boundary terms arise. In formal systems this yields the fourfold structure Q_F = {Provable_F, Refutable_F, True\Provable_F, False\Refutable_F}. Gödel's sentence occupies the boundary term True\Provable_F: it is true, yet not provable inside the system. In this sense, incompleteness is interpreted as the event in which the binary opposition of provability and refutability fails to internalize the full boundary of truth and falsehood. The paper also relates this interpretation to Tarski's undefinability theorem. Gödel shows that truth cannot be reduced to provability within a sufficiently strong consistent system, while Tarski shows that the truth predicate of a sufficiently strong language cannot be fully defined within that same language. Together, these results reveal the boundary between object language and metalanguage. The Superposition Principle reads this boundary as the place where the same self-referential structure appears in opposite forms: as truth and as unprovability. Gödel's sentence is therefore interpreted as a boundary event in which the same self-referential structure is manifested oppositely within a formal field of judgment.

Keywords: Gödel, incompleteness, Superposition Principle, opposition, negation, Tarski, truth, falsehood, provability, self-reference, formal systems

1. Introduction

Gödel's incompleteness theorems are among the central results of modern mathematical logic. The first incompleteness theorem shows that any sufficiently strong and consistent formal system contains a sentence that is true but not provable within that system. In compressed form, the result may be represented as follows:

$$\exists G \ (\mathrm{True}(G) \wedge \neg \mathrm{Prov\_F}(G) \ ).$$

Here F is a formal system strong enough to represent arithmetic, and Prov_F(G) means that the sentence G is provable in F. The standard philosophical summary is that truth and provability do not coincide:

$$\mathrm{Truth} \neq \mathrm{Provability\_F}.$$

This paper accepts the mathematical force of Gödel's theorem and asks a structural question: why does a formal system become incomplete when it attempts to express its own conditions of proof? The answer proposed here is given through the Superposition Principle. A formal system becomes incomplete because the boundary through which truth and falsehood are differentiated cannot be completely owned by the internal rules of proof and refutation.

The thesis may be stated as follows: incompleteness arises when the truth-falsehood boundary opened by a structure of judgment is partially internalized by a formal system, but cannot be fully closed by that system's own proof rules. Gödel's sentence is the site where this boundary becomes explicit.

The interpretation developed here also requires Tarski's undefinability theorem. Gödel shows a limit of provability. Tarski shows a limit of internal truth definition. Together they reveal the boundary between object language and metalanguage. The Superposition Principle provides a language for reading that boundary.

2. Gödel's Sentence and the Boundary of Self-Reference

Let F be a formal system with axioms, rules of inference, and a formal language. For a sentence φ, write Prov_F(φ) for the claim that φ is provable in F. The formalist ideal seeks a close relation between truth and provability: what is provable should be true, and what is true should be provable.

Soundness has the direction:

$$\mathrm{Prov\_F}(\varphi) \rightarrow \mathrm{True}(\varphi).$$

Completeness would require the converse:

$$\mathrm{True}(\varphi) \rightarrow \mathrm{Prov\_F}(\varphi).$$

Gödel's first incompleteness theorem shows that this converse fails for sufficiently strong consistent systems. A Gödel sentence G may be informally expressed as:

G: 'G is not provable in F.'

More formally:

$$G \leftrightarrow \neg \mathrm{Prov\_F}(\ulcorner G \urcorner).$$

The notation $\ulcorner G \urcorner$ denotes the Gödel number of G. Gödel numbering allows syntactic facts about formulas and proofs to be represented arithmetically inside the system. As a result, F can speak, in coded form, about its own proof structure.

This is the crucial boundary. G belongs to the language of F, since it is a sentence constructed within that language:

$$G \in \mathrm{Lang}(F).$$

At the same time, G speaks about provability in F as a whole. In this respect it has a metalinguistic character:

$$G \in \mathrm{Meta}(F).$$

Thus the Gödel sentence lies at the intersection of object language and metalanguage:

$$G \in \mathrm{Lang}(F) \cap \mathrm{Meta}(F).$$

The Superposition Principle reads this intersection as a boundary event. The sentence is internal to the system, yet its truth cannot be exhausted by the system's internal proof relation.

3. Tarski and the Undefinability of Truth

The phrase 'internalization of the truth-falsehood boundary' cannot be grounded in Gödel alone. Gödel's theorem concerns provability. Tarski's theorem concerns truth. Tarski's undefinability theorem states, in broad terms, that a sufficiently expressive formal language cannot define its own truth predicate within that same language.

In compressed form:

$$\mathrm{Truth\_F} \text{ is not fully definable in } \mathrm{Lang}(F).$$

The truth predicate for a language requires a metalanguage. To say which sentences of an object language are true, one must move to a language capable of speaking about that object language. Tarski therefore clarifies the structural separation between object language and metalanguage.

Gödel and Tarski reveal complementary limitations:

$$\text{Gödel: Truth} \neq \text{Provability\_F}.$$

$$\text{Tarski: Truth\_F cannot be fully defined in Lang}(F).$$

Gödel shows that truth cannot be reduced to provability inside a sufficiently strong consistent system. Tarski shows that truth itself cannot be completely defined inside the same language whose truth is at stake. Together, these results expose a boundary that no sufficiently strong formal system can fully absorb into itself.

The Superposition Principle interprets this as the impossibility of completely internalizing the truth-falsehood boundary. The point is not to replace the formal theorems, but to describe the common structure they disclose: the object language calls upon a metalanguage precisely where it attempts to close itself.

4. The First Axiom of the Superposition Principle

The first axiom of the Superposition Principle is the axiom of oppositional-pair identity:

$$\forall A, \exists P\_A = (A/A^*) \text{ such that } A \sim A^* \text{ in } P\_A.$$

The meaning of the axiom is this: A and its opposite A* are not two independently given substances that are later connected. Rather, the oppositional pair P_A comes first, and A and A* are opposite manifestations of that single pair in a given field of judgment or context.

This may be expressed in a simple sentence:

The same thing appears in opposite ways within the field of judgment.

Here sameness does not mean numerical equality at the level of judgment. A and A* are distinct as manifested terms:

$$A \neq A^*.$$

Yet they are structurally related within the same oppositional pair:

$$A \sim A^* \text{ in } P\_A.$$

Applied to truth and falsehood, the axiom gives:

$$P\_T = (T/F).$$

Truth and falsehood are distinct at the truth-value level:

$$T \neq F.$$

Yet they are structurally related within one truth-falsehood pair:

$$T \sim F \text{ in } P\_T.$$

Thus T ~ F does not mean T = F. It means that truth and falsehood are opposite manifestations of one structure of judgment. Classical logic distinguishes the values; the Superposition Principle interprets the structural field in which such distinction becomes possible.

5. The Formal Distinction between T ~ F and T != F

A precise distinction is necessary. Without it, the claim that truth and falsehood are structurally related may be misunderstood as a collapse of classical logic. The paper therefore separates two levels.

At the truth-value level:

$$T \neq F.$$

A sentence cannot, in the same respect and within the same formal context, be both true and false in classical logic. The law of non-contradiction remains in force.

At the structural-pair level:

$$P\_T = (T/F).$$

The notation T ~ F is defined as follows:

$$T \sim F :\equiv \exists P\_T \ (P\_T = (T/F) \wedge T \in P\_T \wedge F \in P\_T \wedge T \neq F).$$

This definition makes clear that T ~ F is not truth-value identity. It is structural-pair relatedness. Truth and falsehood are mutually determined as opposite terms within one judgmental structure.

Accordingly, the paper maintains both claims:

$$T \neq F \text{ at the truth-value level.}$$

$$T \sim F \text{ at the structural-pair level.}$$

This distinction allows the Superposition Principle to interpret classical truth and falsehood without dissolving their difference.

6. The Principle of Superposition of Opposition and Negation

The first axiom explains opposition. Gödel's theorem, however, requires more than simple opposition. It requires the superposition of opposition and negation. A formal system does not merely distinguish truth from falsehood; it also distinguishes provability from unprovability, refutability from unrefutability, and internal judgment from metalinguistic excess.

We therefore introduce the Principle of Superposition of Opposition and Negation:

$$P\_A = (A/A^*) \Rightarrow Q\_A = \{A, A^*, \neg\_C\ A, \neg\_C\ A^*\}.$$

Here A and A* are opposite manifestations of one oppositional pair. The terms ¬_C A and ¬_C A* are context-dependent negative or boundary terms. Negation does not merely erase A. It marks the boundary at which A cannot close itself as A within the context C.

Applied to formal systems, the binary pair of provability and refutability is:

$$P\_F = (\text{Provable}\_F \ / \ \text{Refutable}\_F).$$

But Gödel's theorem shows that this binary structure does not close the whole field of judgment. Self-reference opens a fourfold structure:

$$Q\_F = \{Provable\_F, Refutable\_F, True\setminus Provable\_F, False\setminus Refutable\_F\}.$$

The four terms may be read as follows: provable sentences; refutable sentences; true but unprovable sentences; and false but unrefutable sentences. Gödel's sentence occupies the boundary term:

$$G \in True\setminus Provable\_F.$$

That is:

$$True(G) \wedge \neg Prov\_F(G).$$

Thus G is not merely a true sentence. It is a boundary term in which truth and unprovability overlap. Incompleteness appears because the binary opposition of provability and refutability cannot fully internalize the boundary of truth and falsehood. The system attempts to close judgment within P_F, but self-reference opens Q_F.

In compressed interpretive form:

$$Incompleteness(F) = Q\_F \setminus P\_F.$$

This expression means that incompleteness is the remainder opened when the fourfold structure of truth, falsehood, provability, and unprovability exceeds the binary closure of provability and refutability.

7. Gödel's Sentence as Opposite Manifestation of the Same Self-Referential Structure

The Gödel sentence has the form:

$$G \leftrightarrow \neg Prov\_F(\ulcorner G \urcorner).$$

If G is provable in F, then what G says is false:

$$Prov\_F(G) \rightarrow False(G).$$

If G is not provable in F, then what G says is true:

$$\neg Prov\_F(G) \rightarrow True(G).$$

Thus the same self-referential structure appears in opposite ways within the formal field of judgment. Under the condition of provability, it appears as false. Under the condition of unprovability, it appears as true.

This is precisely the pattern expressed by the first axiom of the Superposition Principle. The same underlying structure is not first divided into two independent entities. Rather, it is manifested oppositely through the conditions of judgment.

Let P_G denote the self-referential oppositional structure of the Gödel sentence:

$$P\_G = (G/G^*).$$

Here G is the direction 'I am not provable,' while G* is the opposite direction 'I am provable.' Within the formal context F, this structure differentiates as follows:

$$P\_G \xrightarrow{[Prov\_F(G)]} False(G).$$

$$P\_G \xrightarrow{[\neg Prov\_F(G)]} True(G).$$

The Gödel sentence is therefore not simply assigned the label 'true' from outside. It is the site at which the same self-referential structure is manifested as truth through unprovability and as falsehood through provability.

8. Why Formal Systems Become Incomplete

We can now answer the guiding question of the paper: why is a sufficiently strong formal system incomplete? The answer is that the formal system attempts to close the field of judgment through its internal rules of proof and refutation, but the truth-falsehood boundary exceeds that closure when self-reference appears.

The system seeks to internalize truth as provability and falsehood as refutability:

$$T \rightarrow \text{Provable\_F}.$$

$$F \rightarrow \text{Refutable\_F}.$$

This would mean that the truth-falsehood boundary is fully captured by the provability-refutability pair P_F. But Gödel's construction shows that, once the system speaks about its own provability, the object language and the metalanguage overlap:

$$\text{Lang}(F) \otimes \text{Meta}(F).$$

At this boundary the Gödel sentence G is generated:

$$G \in \text{Lang}(F) \cap \text{Meta}(F).$$

The sentence is formed within the language of the system, yet its truth is not reducible to the system's internal proof relation:

$$\text{True}(G) \wedge \neg\text{Prov\_F}(G).$$

Therefore the system is incomplete not merely because an isolated sentence has escaped proof. It is incomplete because the fourfold structure Q_F remains larger than the binary closure P_F. The true-but-unprovable boundary term cannot be absorbed by the internal rules of proof.

The central interpretive thesis is therefore:

A formal system becomes incomplete because the boundary through which the truth-falsehood pair is differentiated cannot be fully internalized by the system's own rules of provability and refutability.

9. The Second Incompleteness Theorem

The second incompleteness theorem states that a sufficiently strong consistent formal system cannot prove its own consistency. If Con(F) expresses the consistency of F, then:

$$\text{Con}(F) \rightarrow \neg\text{Prov\_F}(\text{Con}(F)).$$

Consistency may be represented as the non-existence of a sentence φ such that both φ and its negation are provable:

$$\text{Con}(F)\text{: } \neg\exists\varphi \text{ (Prov\_F}(\varphi) \wedge \text{Prov\_F}(\neg\varphi) \text{)}.$$

From the standpoint of the Superposition Principle, consistency is the claim that the truth-falsehood boundary has not collapsed inside the system. But the system cannot completely guarantee that boundary from within itself.

Thus the second incompleteness theorem deepens the boundary analysis. The first theorem shows the emergence of a boundary sentence G. The second theorem shows that the system cannot fully certify, from within, that its own boundary of non-contradiction remains intact.

In this sense:

$$\partial\text{\_C F is not fully contained in Prov\_F}.$$

The boundary of the system cannot be completely owned by the system.

10. Conclusion

This paper has interpreted Gödel's incompleteness theorems through the Superposition Principle. The central claim is that incompleteness can be read as the impossibility of completely internalizing the boundary through which truth and falsehood are differentiated within a formal system.

The first axiom of the Superposition Principle states that an item and its opposite are not two independently given substances, but opposite manifestations of a single pair. Applied to truth and falsehood, this yields P_T = (T/F). Truth and falsehood are distinct at the truth-value level, T != F, yet structurally related within one pair, T ~ F in P_T.

The Principle of Superposition of Opposition and Negation extends this binary relation into a fourfold structure. In formal systems, the system attempts to close judgment within the pair P_F = (Provable_F/Refutable_F). Gödel's construction opens Q_F = {Provable_F, Refutable_F, True\Provable_F, False\Refutable_F}. Gödel's sentence belongs to the boundary term True\Provable_F.

The sentence G is internal to the formal language and yet requires a metalinguistic perspective. It therefore appears at the boundary Lang(F) ∩ Meta(F). Tarski's undefinability theorem clarifies the same boundary from the side of truth: the truth predicate of a sufficiently strong language cannot be fully defined within that same language.

The final thesis may be stated as follows:

Gödelian incompleteness is a boundary event in which the same self-referential structure is manifested oppositely as truth and unprovability, and in which the binary opposition of proof and refutation fails to internalize the full truth-falsehood boundary.

In this interpretation, the Superposition Principle does not erase the difference between truth and falsehood. It explains the structural field in which that difference emerges, and it shows why a formal system that attempts to close that field from within gives rise to incompleteness.

References

Boolos, George S., John P. Burgess, and Richard C. Jeffrey. Computability and Logic. Cambridge: Cambridge University Press, 2007.

Feferman, Solomon. 'Arithmetization of Metamathematics in a General Setting.' Fundamenta Mathematicae 49, 1960.

Gödel, Kurt. 'Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I.' Monatshefte für Mathematik und Physik 38, 1931.

Hilbert, David. 'Die Grundlagen der Mathematik.' Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg 6, 1928.

Kleene, Stephen C. Introduction to Metamathematics. Amsterdam: North-Holland, 1952.

Nagel, Ernest, and James R. Newman. Gödel's Proof. New York: New York University Press, 1958.

Smith, Peter. An Introduction to Gödel's Theorems. Cambridge: Cambridge University Press, 2007.

Tarski, Alfred. 'The Concept of Truth in Formalized Languages.' In Logic, Semantics, Metamathematics. Oxford: Clarendon Press, 1956.