QUESTIONING GÖDEL'S ONTOLOGICAL PROOF:
IS TRUTH POSITIVE?

GREGOR DAMSCHEN

Martin Luther University of Halle-Wittenberg

Abstract. In his "Ontological proof", Kurt Gödel introduces the notion of a second-order value property, the positive property *P*. The second axiom of the proof states that for any property *φ*: If *φ* is positive, its negation is not positive, and vice versa. I put forward that this concept of positiveness leads into a paradox when we apply it to the following self-reflexive sentences: (A) The truth value of A is not positive; (B) The truth value of B is positive. Given axiom 2, sentences A and B paradoxically cannot be both true or both false, and it is also impossible that one of the sentences is true whereas the other is false.

If Oscar Wilde is right and the way of paradoxes is the way to truth, then the way of the paradoxes of truth would (paradoxically) also be the way to truth. Paradoxes of truth, like the liar and the strengthened liar, have always been the source of substantial reflections on the concept of truth. Every new discovery in the field of the paradoxes of truth can, therefore, increase our understanding of truth. For this reason, I would like to present a new problem, perhaps even a new paradox of truth, which arises out of the question whether *truth is positive*. We shall see that Kurt Gödel's positiveness property (in particular axiom 2 of the Gödelian proof of the existence of God) together with two self-referential sentences containing statements about the positiveness of truth values leads to a paradox of positive truth.

I. GÖDEL-POSITIVENESS

A formal system of positiveness, which is to be our starting point here, has been sketched out by Kurt Gödel in his "Ontological proof" of the

existence of God (Gödel 1995, 403-4). This modal logical argument dated from February 10, 1970 was familiar to logicians and philosophers since the seventies of the previous century, but was not published until 1987 (Sobel 1987, 256-7). At the centre of the proof is the assumption of the so-called positive property.

The basic concept of positiveness P that Gödel uses in his proof is explained in the following way: Positiveness P is a second-order property, which is applied to first-order properties φ. In the first line of the ontological proof, positiveness is therefore introduced as the property $P(\varphi)$: "φ is positive (or $\varphi \in P$)" (Gödel 1995, 403). Secondly, 'P' is a value predicate, positiveness P is a value property: "Positive means positive in the moral aesthetic sense (independently of the accidental structure of the world)" (Gödel 1995, 404). There is a notebook entry by Gödel that provides us with more information about what we are to understand by "the moral aesthetic sense" of a positive property: "It is possible to interpret the positive as perfective; that is, "purely good," that is, such as implies no negation of "purely good."" (Gödel 1995, 435 = notebook "Phil XIV", p. 105)[1]. And so, thirdly, positiveness is not a value property that only receives its value in relation to something else; in fact, it possesses its value absolutely. In demanding that positiveness is to be understood as "pure good" Gödel explicitly distances himself from the assumption that positive is in the first place that which is good in some respect: "The interpretation of "positive property" as "good" (that is, as one with positive value) is impossible, because the greatest advantage + the smallest disadvantage is negative" (Gödel 1995, 435 = notebook "Phil XIV", p. 105)[2]. Fourthly, Gödel's positiveness is orientated towards the concept of perfections – whereby perfections are properties that contradict neither themselves nor any other perfection: "It [sc. positive] may also mean pure "attribution" as opposed to "privation" (or *containing*

---

[1] The ontological proof itself is written in English. Gödel's notebook entries, which are written in German, were translated into English by Robert M. Adams (cf. Gödel 1995, 429). The German original reads: "Es ist möglich, die positive als perfectiv zu interpretieren, d. h., "rein gut", d. h., solche, welche keine Negation von "rein gut" impliziert." (Gödel 1995, 434).

[2] German original: "Die Interpretation von "positiver Eigenschaft" als "guter" (d. h., einer mit positivem Wert) ist unmöglich, weil der größte Vorteil + dem kleinsten Nachteil negativ ist." (Gödel 1995, 434).

privation)" (Gödel 1995, 404), and "a property is a perfective if and only if it implies no negation of a perfective" (Gödel 1995, 435 = notebook "Phil XIV", p. 106)[3]. Linked with these clarifications of the content of Gödelian positiveness is a final fifth assumption: If a first-order property is positive, then its positiveness is invariant with respect to the bearer of the first-order property ("independently of the accidental structure of the world", Gödel 1995, 404). In summary, Gödel-positiveness is at least a second-order perfect invariant value property.[4]

II. ONE AXIOM OF POSITIVENESS
AND TWO SENTENCES ON TRUTH

Let us now take a widely held assumption concerning truth:

Assumption 1   Truth is a first-order property of bearers of truth values.

If truth is a first-order property and if positiveness is a second-order property, the question arises whether truth itself is positive or not. This question of positive truth creates problems when we employ two other assumptions. The first problem maker is Axiom 2 of Gödel's proof, which states that for each and every property φ the following is true: either the property itself or its complement is positive (Gödel 1995, 403).[5]

Axiom 2 $P(\neg\varphi) \leftrightarrow \neg P(\varphi)$[6]
             In every pair consisting of a property and its negation *one and only one* property is positive.

---

[3] German original: "Eine Eigenschaft ist eine Perfective, dann und nur dann wenn sie keine Negation einer Perfectiven impliziert" (Gödel 1995, 434).

[4] In the following, "positiveness" and "positive" always stand for "Gödel-positiveness" and "Gödel-positive."

[5] In Sobel's and other publications, which follow notes on Gödel's proof in Dana Scott's hand, this axiom is Axiom 1 (Sobel 1987, 242; 257; Sobel 2004, 119).

[6] Gödel originally uses the following notation: $P(\varphi) \vee P(\sim\varphi)$, where "$\vee$" is an exclusive "or". This is equivalent to Scott's notation $P(\neg\varphi) \leftrightarrow \neg P(\varphi)$ (Sobel 1987, 257) which is short for $\Box \forall\varphi[P(\neg\varphi) \leftrightarrow \neg P(\varphi)]$ (see Sobel 2004, 119).

This axiom can easily be divided up into two axioms, each of which must then be treated separately:

Axiom 2.1 $\neg P(\varphi) \rightarrow P(\neg\varphi)$[7]
             At least one member of every pair consisting of a property and its negation is positive.

Axiom 2.2 $P(\neg\varphi) \rightarrow \neg P(\varphi)$[8]
             At the most one member of every pair consisting of a property and its negation is positive.

The second group of problem makers is the following group of two self-referential sentences concerning truth values:

(A) The truth value of A is not positive.
(B) The truth value of B is positive.

The self-referential sentences A and B are not trivial, not paradoxical and not meaningless. They are not trivial because they are neither analytical nor can they be derived from analytical sentences. Their truth value cannot be determined immediately. They are not paradoxical, since it is not the case that they are true if and only if they are false. Furthermore, they are not meaningless, since they are syntactically well formed and the concepts used in them, when taken together, produce a thought that can be considered as true or false. The self-referentiality of the two sentences is on its own no reason to consider them meaningless, since such sentences as "This sentence contains five words" seem to be capable of truth.

III. ARE TRUTH VALUES POSITIVE?

Accepting these assumptions one may ask what truth value the sentences A and B have. For A, there are two possibilities.

---

[7] Short for $\Box\,\forall\varphi[\neg P(\varphi) \rightarrow P(\neg\varphi)]$.
[8] Short for $\Box\,\forall\varphi[P(\neg\varphi) \rightarrow \neg P(\varphi)]$.

If A is true, the truth value of A is not positive. If the truth value of A is not positive, then what A expresses is the case and A is true. Therefore, A is true if and only if the truth value of A is not positive.

If A is not true, then what A expresses is not the case and the truth value of A is positive. If the truth value of A is positive, then what A expresses is not the case and A is not true. Therefore, A is not true if and only if the truth value of A is positive.

If A is true, the truth value of A cannot at the same time be not true, since A would then be true and not true at the same time and this would be a contradiction. The same holds good for the case in which A is not true. Here, too, it is the case that when A is not true its truth value is also not true. The result, therefore, is that the truth of A is not a positive property and the non-truth of A is a positive property:

1. The non-truth of A is positive.
2. The truth of A is not positive.

Applying the axiom 2.1 we get for sentence 2 the result:

3. If the truth of A is not positive, then the non-truth of A is positive.

Taking sentence 2 and 3 together we get *per modus ponens* the result:

4. The non-truth of A is positive.

In the case of sentence B one accordingly gets different results:

5. The non-truth of B is not positive.
6. The truth of B is positive.

Applying axiom 2.1 to sentence 5 produces the result:

7. The truth of B is positive.

If we accept axiom 2.1, we have the result that, both on the assumption of the truth of sentence A and on the assumption of the non-truth of sentence A, the non-truth of A is positive. In the case of B the situation is the other way round. Both on the assumption of the truth of B and on the assumption of the non-truth of B, the truth of B is positive.

IV. POSITIVENESS IS INVARIANT

If positive properties are perfections and positiveness relates to the perfection of the property, then the perfection of the property is not affected by the fact that it may possibly be exemplified by various individuals. If, for example, pure beauty is positive, then it is a matter of indifference with respect to the positiveness of pure beauty who or what exhibits pure beauty.[9]

Let us now take, in addition, the following axiom of positiveness invariance (PosInv):

Axiom PosInv   For all x, for all φ: If something x possesses the property φ and φ is positive, then the positiveness of φ is invariant with respect to the bearer x of φ.

If we apply the axiom of positiveness invariance to sentence 4, we get the result:

8.   If the non-truth of A is positive, then non-truth is positive.

Taking sentence 4 and 8 together we get *per modus ponens* the result:

9.   Non-truth is positive. $P(\neg T)$

If we apply the axiom of positiveness invariance to sentence 7, we get the result:

10.   If the truth of B is positive, then truth is positive.

Finally, taking sentence 7 and 10 together we get the result:

11.   Truth is positive. $P(T)$

---

[9] See also Gödel's axiom 3 (= Sobel 1987 axiom 4): $P(\varphi) \rightarrow \Box P(\varphi)$ (Gödel 1995, 403).

V. THE PARADOX OF POSITIVE TRUTH

There are precisely four ways in which the meaningful sentences A and B can relate to one another logically: both sentences are true, both are not true, or, in each case, one is true and the other is not true. A simple formal proof shows that each of the four instances leads to a formal contradiction if we add axiom 2.2.

1. $T(A) \wedge T(B)$: $P(\neg T) \wedge P(T)$

$P(\neg T) \wedge P(T)$ + axiom 2.2   leads to *contradiction:*
$\neg P(\neg T) \wedge P(\neg T)$

Proof:

1. $P(\neg T) \wedge P(T)$	
2. $\forall \varphi (P(\neg \varphi) \rightarrow \neg P(\varphi))$	axiom 2.2
3. $\forall \varphi (P(\varphi) \rightarrow \neg P(\neg \varphi))$	from 2, contrapositive of axiom 2.2
4. $P(T) \rightarrow \neg P(\neg T)$	$\forall$-elimination, T/φ in 3
5. $P(T)$	from 1, simplification
6. $\neg P(\neg T)$	from 4 und 5, modus ponens
7. $P(\neg T)$	from 1, simplification
Therefore: 8. $\neg P(\neg T) \wedge P(\neg T)$	from 6 and 7, introduction of conjunction

Q.E.D.

2. $\neg T(A) \wedge \neg T(B)$: $P(\neg T) \wedge P(T)$

$P(\neg T) \wedge P(T)$ + axiom 2.2   leads to *contradiction:*
$\neg P(\neg T) \wedge P(\neg T)$

Proof as in 1.

3. $T(A) \wedge \neg T(B)$: $P(\neg T) \wedge P(T)$

$P(\neg T) \wedge P(T)$ + axiom 2.2   leads to *contradiction:*
$\neg P(\neg T) \wedge P(\neg T)$

Proof as in 1.

4. $\neg T(A) \wedge T(B)$: $P(\neg T) \wedge P(T)$

   $\qquad\qquad P(\neg T) \wedge P(T)$ + axiom 2.2   leads to *contradiction:*

   $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \neg P(\neg T) \wedge P(\neg T)$

   Proof as in 1.

VI. RESULTS

The statements A and B concerning the positiveness of their truth values produce, on the assumption of Gödel's axioms 2.1, 2.2 and the axiom of positiveness invariance, a paradox. The individual sentences A and B taken together can neither be true nor not true; nor can, in each case, the one be true and the other not true. It is clear that each of the four possibilities produces a formal contradiction. We assumed, however, that the sentences A and B are not trivial, not paradoxical and not meaningless.

This result is valid not only for Gödel's positive property P but for any second-order property ψ that is integrated into a formal system that, in turn, contains axioms that are at least structurally analogous to the axioms 2.1, 2.2 and to the axiom of positiveness invariance. The puzzle, which has to be solved, arises when one takes in addition self-referential statements like A and B concerning the positiveness (or some ψ-ness) of the truth values of A and B. Hence, it seems that the problem is not caused by Gödel's positiveness property alone but also – as in the other familiar paradoxes of truth – by the predicate of truth and by the self-referentiality of the sentences. How the problem is produced by the joint effect of predicates of positiveness and truth is a question that still remains to be solved.[10]

BIBLIOGRAPHY

Gödel, Kurt. 1995. "Ontological proof". In his *Collected Works*, vol. III, ed. S. Feferman et al., 403-4. Oxford U.P. (plus "Appendix B: Texts relating to the ontological proof", 429-37).

Sobel, Jordan Howard. 1987. "Gödel's Ontological Proof". In *On Being and Saying. Essays for Richard Cartwright*, ed. J. J. Thomson, 241-61. Cambridge, MA.

Sobel, Jordan Howard. 2004. "Kurt Gödel's *Ontologischer Beweis*". In his *Logic and Theism. Arguments For and Against Beliefs in God*, 115-167. Cambridge U.P.