Introductory note to **1970*

1. Background

Gödel showed his **1970* to Dana Scott, and discussed it with him, in February 1970. Gödel was very concerned about his health at that time, feared that his death was near, and evidently wished to insure that this proof would not perish with him. Later in 1970, however, he apparently told Oskar Morgenstern that though he was "satisfied" with the proof, he hesitated to publish it, for fear it would be thought "that he actually believes in God, whereas he is only engaged in a logical investigation (that is, in showing that such a proof with classical assumptions [completeness, etc.], correspondingly axiomatized, is possible)."[a]

Scott made notes on the proof and presented a version of the argument to his seminar on logical entailment at Princeton University in the fall of 1970. Through this presentation and the recollections and notes of those who attended the seminar, Gödel's ontological proof has become fairly widely known. Discussion of the proof, thus far, has been based largely on Scott's version of it (*Scott 1987*), which differs somewhat in form from Gödel's own memorandum. The latter is published here—though not for the first time; like Scott's version, it was published as an appendix to *Sobel 1987*, pages 256–7.

Gödel had devised his ontological proof some time before 1970. Other, presumably earlier, versions of it have been found among his papers. A sheet of paper headed "Ontological Proof" (in German), and dated, in Gödel's own hand, "ca. 1941", contains some but not all of the ideas of the proof. Extensive preparatory material is contained in the philosophical notebook "Phil XIV". The first page of this notebook bears a notation indicating that it was written during the period "Ca. July 1946–May 1955". The last page of the notebook contains the note "Asbury Park 1954 p. 100 ff.", which presumably applies to the pages (103–109) pertaining to the ontological proof. Other documents, including letters, indicate that Gödel intended to leave Princeton for the shore 9 August 1954, was vacationing in Asbury Park on 25 August 1954, and was probably back in Princeton by 3 October 1954. We may reasonably assume, then, that the notebook pages on the ontological proof were written in the late summer and early fall of 1954 and were completed at any rate

[a]Morgenstern's diary for 29 August 1970, Box 15 of the Oskar Morgenstern Papers, quoted by courtesy of the Special Collections Department, Duke University Library, Durham, North Carolina. I am indebted to John Dawson for noticing and communicating this item.

by May 1955.[b] Relevant excerpts from the notebook, and two of the (presumably earlier) loose sheets headed "Ontological Proof", including the one dated "ca. 1941", are published in Appendix B to this volume.

Among the historic sponsors of the ontological argument, it is not to Anselm or Descartes but to Leibniz that the parentage of Gödel's proof belongs, as scholars interested in the proof have long recognized (see, e.g., *Sobel 1987*, page 241). The study of Leibniz is known to have been a major intellectual preoccupation for Gödel during the 1930's (*Menger 1981*, §§8, 12), and especially during 1943–46 (*Wang 1987*, pages 19, 21, 27). Little discussion of Leibniz's treatment of the ontological argument as such has been found in Gödel's papers, but he must have known two things about it:

1. Leibniz held that Descartes's ontological proof is incomplete. It does succeed in proving the conditional proposition that *if* God's existence is so much as possible, then God actually (and indeed necessarily) exists. But it assumes without proof that God's existence is possible; and that, Leibniz argues, must be proved in order to complete the demonstration. Leibniz says this in many places in his writings, some of them so familiar to students of Leibniz that Gödel can safely be assumed to have known them (e.g., *Leibniz 1969*, pages 292–3). In January 1678 Leibniz wrote down an elaborate and interesting proof of the conditional proposition (*Leibniz 1923–*, II, i, 390–1), but I have seen no specific evidence that Gödel was familiar with that text.

2. Leibniz also held that the ontological proof can be completed by proving the possibility of God's existence. His main attempt to accomplish this is based on a conception of God as *Ens perfectissimum*, a being whose attributes are all the perfections, where perfections are identified with simple, purely positive qualities, and where a purely positive quality cannot be limited and therefore cannot be an inferior degree of any quality. Leibniz argues that purely positive qualities must all be consistent with each other, so that no inconsistency can arise from the conception of an *Ens perfectissimum*, which must therefore be a possible being. This argument is most fully developed in texts Leibniz wrote in 1676 (*Leibniz 1923–*, VI, iii, 395–6, 571–79). It recurs with almost cryptic brevity at the end of his life in §45 of the famous "Monadology" of 1714, but it has become known mainly through one of the 1676 texts, "That an *Ens Perfectissimum* Exists", which has been generally accessible, both in Latin and in English translation, since the end of the nineteenth century (*Leibniz 1923–*, VI, iii, 578–9 = *1969*, pages 167–8).

[b]I am indebted to John and Cheryl Dawson for the information on dating cited here.

This text at least, and the "Monadology", were surely known to Gödel, whose ontological proof is built around an idea of positive properties.

Gödel's treatment of the ontological proof resembles Leibniz's on both of these points. The first point will be the subject of §2 of this introduction. The second will occupy us in §§3–4.

2. If possible, then actual

Gödel resembles Leibniz in making the ontological proof proceed by way of the conditional thesis that if the divine existence is so much as possible, then it is actual, and indeed necessary. In his **1970*, this thesis occurs as the line

$$M(\exists x)G(x) \supset N(\exists y)G(y),$$

which I shall be calling (iii). (I follow Gödel in using M and N as possibility and necessity operators, respectively.) As noted above, however, Gödel shows no clear influence of Leibniz's fullest argument for the thesis, which turns on a rather different conception of "essence" from Gödel's.

The grounds Gödel gives for the conditional thesis show more affinity with a type of "ontological argument" based on modern modal logic which has gained currency in the last thirty years. Charles Hartshorne published a proof of this type in his *1962* (pages 50–53), and subsequent discussion has established its logical properties quite clearly (see *Lewis 1970*; *Adams 1971*; *Plantinga 1974*, pages 196–221). In a presentation approximating Hartshorne's, the first part of the proof has the following steps, which are found also in Gödel's proof:

(i) $N[(\exists x)G(x) \supset N(\exists y)G(y)]$
(ii) $M(\exists x)G(x) \supset MN(\exists y)G(y)$
(iii) $M(\exists x)G(x) \supset N(\exists y)G(y)$.

Step (i) is the necessitation of the line immediately following, and inferred from, the theorem following Axiom 4 in Gödel's **1970*. Since Gödel takes this line to be entailed by a theorem, he may be presumed to accept its necessitation.[c] Step (i) is the thesis that it is necessary that if God exists at all, God exists necessarily—or, more briefly, that it is impossible for God to exist contingently. Some philosophers have thought that the concept of God is a concept of a Necessary Being, and that (i) follows straightforwardly from the concept of God (cf. *Hartshorne 1962*, page 41; *Findlay 1965*). Gödel gives a more complicated derivation of (i), which hinges on the claim (Axiom 4) that necessary existence is a positive property. Since he has made it true by definition, and hence necessarily true, that God (if God exists) has all positive properties, and since (by Axiom 3) any property that is positive is necessarily positive, it follows that God (if God exists) has necessary existence. That is, (i) follows from these assumptions. This strategy for proving (i) is obviously akin to the attempts that have been made, in the history of the ontological argument, to derive something equivalent to (i) from the claim that necessary existence is a "perfection", it being assumed that God, by definition, possesses all perfections.[d]

Step (ii) is inferred from the line that corresponds to (i) in Gödel's proof, and does indeed follow from (i) by the principle

(iv) $N(p \supset q) \supset (Mp \supset Mq)$,

which would be an axiom or theorem in any system of modal logic that would be likely to be used in this context. The inference from (ii) to (iii), on which Gödel also relies, depends on a more controversial principle,

(v) $MNp \supset Np$,

which is a form of the characteristic axiom of S5, the most powerful of the standard systems of modal propositional logic. One of the firm results of recent studies of modal versions of the ontological argument is that (iii) does follow from (i) in S5.[e] Whether it is appropriate in this context to rely on S5, and particularly on (v), is certainly open to question, but several philosophers have believed that it is appropriate.[f] Gödel must apparently be counted among them, though he may have had some reservation on this point.[g]

There is no evidence that Gödel was influenced by the recent work of others on modal ontological proofs. The derivation of (iii) from (i) in S5 had been published by Hartshorne in 1962 and was attracting the attention of other students when Gödel showed his ontological proof to Scott in 1970. But, as already noted, Gödel had developed his proof

[c]The presumption is made explicit in Sobel's reconstruction of the proof (*1987*, pp. 247–8).

[d]See *Anselm 1974*, pp. 94–5 (*Proslogion*, chapter 3); *Malcolm 1960*, p. 46.

[e]See *Hartshorne 1962*, pp. 39–40, 51–53; *Plantinga 1974*, pp. 213–17. In one sense, S5 is more than is needed. A similar modal ontological proof can be constructed in the somewhat weaker modal system sometimes called "Brouwerian", in which (v) is replaced by the axiom $p \supset NMp$ (*Adams 1971*, pp. 40–48). But there is no strong reason for thinking the Brouwerian system more acceptable than S5 in this context.

[f]*Hartshorne 1962*, pp. 39–40, 51–53; *Adams 1971*, pp. 42, 45–6; *Plantinga 1974*, p. 215; *Sobel 1987*, p. 246.

[g]Morton White (in personal correspondence) reports that Gödel expressed "reservations about his ontological proof because of his doubt about using some principle in modal logic", but that Gödel did not specifically mention S5 or its characteristic axiom. So far as I am aware, this is the only point in the proof about which Gödel is known to have expressed a reservation.

some years earlier. His notebook entries on the proof, from 1954 or 1955, do not articulate the modal logic used in the proof, but there is no reason to doubt that he was already consciously relying on (v) or on something equivalent to it. One of Gödel's early sketches for the ontological proof, dating perhaps from the 1940's, ends with an inference precisely from (ii) to (iii), in which he must be relying implicitly on (v) as a principle.[h] Gödel may in fact have been the first student of modern logic to see that this principle could be used to prove that "if the concept of necessary existence is consistent, then there are things to which it applies", as he put it in that early sketch.

One problem about the logical apparatus of Gödel's **1970* should be noted. Definition 2 fails to imply that every essence of x must be true of x. It implies, indeed, that if there is any property that is necessarily false of everything, it is an essence of x. Then from the definition of "$E(x)$", with the assumption that there is a property that is necessarily false of everything (an assumption that Gödel seems to make in **1970*, since he treats "$x \neq x$" as expressing a [negative] property), we could further infer that "$E(x)$" is not true of anything. The latter conclusion is obviously contrary to Gödel's intent in the proof. Moreover, the claim in footnote 3, that "any two essences of x are *necessarily equivalent*", also seems to presuppose that every essence of x must be true of x. Scott (*1987*, page 258) doubtless represents Gödel's intention correctly when he adds "$\varphi(x)$" as a conjunct to the right side of the definition of "φ Ess x".[i] It is interesting that the page on which Gödel wrote the early sketch of his ontological proof mentioned in the previous paragraph ends with a note in which Gödel proposes a definition of essence whose right side is like that of Definition 2 of **1970* except that "$\varphi(x)$" is added as a conjunct, so that the definition does imply that every essence of x is true of x.[j]

3. Leibniz's possibility proof

Accepting the conditional thesis that if God's existence is possible, then God exists, one needs only the further premise that God's existence is possible in order to detach the consequent and infer by modus

[h]This sketch is printed in Appendix B to this volume. Of the two such documents reproduced there, it is the one not dated by Gödel.

[i]Scott is followed in this by Sobel (*1987*, page 244) and Anderson (*1990*, page 292). I assume that a quantifier, elsewhere in the right side of Scott's definition of "φ Ess x", printed in Sobel's appendix as "$\forall x$", should be "$\forall \psi$".

[j]Most of the observations in this paragraph are due to Charles Parsons.

ponens that God actually exists. But how to justify the possibility premise? Possibility is often assumed rather easily, but should not be in this case, for at least two reasons. One reason, emphasized by Leibniz, is that the concept of God is the concept of a sort of maximum (a maximum of perfection), and the concept of a maximum can seem innocent at first glance, while representing something really impossible (e.g., "the largest number"; see *Leibniz 1969*, page 211). Another reason, not noted by Leibniz but prominent in recent discussion of modal ontological arguments, is that at the point in such an argument at which a possibility premise is required, it is typically supposed to have been proved that the existence of God is either impossible or necessary ("$M(\exists x)G(x) \supset N(\exists y)G(y)$" in Gödel's proof). In this context, assuming the possibility of God's existence commits one quite directly to the impossibility of God's nonexistence. But why shouldn't the possibility of God's nonexistence be assumed as easily as the possibility of God's existence? (Cf. *Adams 1988*.) So it would be important, in completing a modal ontological proof, to give a *proof* that God's existence is possible.

Leibniz's attempt to accomplish this begins with a conception of God as a being that possesses all perfections. "A perfection", he says, is what he calls "every simple quality that is positive and absolute, or ⟦*seu* = that is⟧ that expresses without any limits whatever it expresses."[k] Three points about this definition claim our attention. (1) Perfections are *qualities*. What is meant here may not be precisely the Aristotelian category of "quality", but it is surely something narrower than we might mean by "property". For instance, it presumably does not include relations. The divine nature is constituted by *internal* properties. (2) The *simplicity* of the perfections plays a part in the best-known formulation of Leibniz's possibility proof, excluding any analysis of them (*Leibniz 1969*, page 167). But this is superfluous, as Leibniz recognized (*1923–*, VI, iii, 572). Pure positiveness is the only characteristic of the perfections that is really needed for the proof, and the only one that appears in the brief version of the proof in the "Monadology" (§45). (3) The final clause of the definition indicates that "absolute" is being used to mean *unlimited*, not qualified by any limitation. And limitation is understood here as a partial negation. "Absolute" is therefore an intensification of "positive": a perfection is a *purely* positive quality, a quality that involves no negation at all. What sort of involvement of negation is excluded will become clear as we examine the strategy of Leibniz's argument.

Leibniz argues that all *simple* positive qualities are mutually compatible, on the ground that if they were not, "one would express the

[k]I give my own translation from *Leibniz 1923–*, VI, iii, 578–9. An English translation of the whole text is found in *Leibniz 1969*, pp. 167–8.

exclusion of the other, and so one of them would be the negative of the other, which is contrary to the hypothesis, for we assumed that they are all affirmative." He argues further that it follows that any conjunction of purely positive qualities is possible, "for if individual ⟦attributes⟧ are thus compatible, pluralities will be too, and therefore also composites" (*1923–*, VI, iii, 572). His argument for possibility depends on the exclusion of negation from the construction of any purely positive quality. It seems to presuppose a conception according to which a purely positive quality must either be a simple positive quality or, if complex, must be constructible from simple positive qualities without the aid of negation.[1] That is the sense in which, for Leibniz, a purely positive quality cannot involve negation.

Leibniz assumes that the only way in which a conjunction of qualities could be impossible is by having, when fully analyzed, two conjuncts, of which one is formally the negation of the other. But a conjunction of purely positive qualities cannot be impossible in this way. For it cannot, when fully analyzed, have any conjunct that is formally the negation of anything.

Since perfections are purely positive qualities, Leibniz infers that the conjunction of all perfections cannot be impossible, and therefore is possible. Treating the possibility of a conjunction of qualities as equivalent to the possibility of the existence of a being possessing all those qualities, he infers that the existence of a being possessing all perfections is possible. And since such a being would satisfy his definition of God, he infers that the existence of God is possible.

Two possible difficulties for this argument may be noted here.

(1) One might question the assumption that the only way in which a conjunction of qualities can be impossible is by a contradiction involving formal negation, occurring between the qualities or arising in their analysis into a conjunction of simpler qualities. There is ample basis in Leibniz's writings for ascribing to him such a formalistic conception of impossibility; but it is not obvious that even his own statements and arguments are always in keeping with it. And some philosophers, including Descartes (*1985*, pages 45–6), for example, have maintained that simple properties can necessarily exclude each other without either of them being analyzable at all, and without either being the formal negation of the other.

(2) A proof that God's existence is possible will not satisfy the needs of a modal ontological argument unless the God whose existence is proved possible is one that must exist necessarily if at all. But why couldn't the

[1]It may even be that he thought the construction must involve no other logical operation besides conjunction.

conjunction of all perfections be exemplified contingently? The obvious move for Leibniz to attempt in response to this question is to hold that necessary existence is one of the perfections. And in the best-known version of his possibility proof he does say at least that *existence* is one of the perfections (*Leibniz 1969*, page 167). But this response is attended with problems. One which soon occurred to Leibniz himself is that it may be doubted whether existence is a *quality*, as perfections must be (*Leibniz 1923–*, II, i, 313); and presumably this doubt would apply to necessary existence as well.

4. Gödel's possibility proof

Gödel's **1970* contains a strategy for proving possibility that differs from Leibniz's in ways that may help Gödel to deal with both of these difficulties, but that may also bring compensating disadvantages in their train. This is because Gödel's **1970* uses a conception of a positive property that is quite different from Leibniz's conception of a perfection. Two differences may be noted here, having to do with the notions of properties and of positiveness, respectively.

1. Gödel's **1970* speaks of the entities in the domain of the predicate variable φ simply as "properties". This category seems not to be restricted to what Leibniz would count as qualities. Gödel's definitions of G and E, and his syntactical treatment of them and of $x = x$ and $x \neq x$, suggest that he was pretty generally willing here to postulate properties corresponding to propositional functions of a single individual variable. Perhaps Gödel would restrict the applicability of his notion of properties more narrowly than this suggests, but no such restriction is found in the text; in particular, nothing excludes relational properties corresponding to propositional functions of several variables.

This certainly makes it easier for Gödel to defend the thesis (his Axiom 4) that necessary existence is a positive property, which he uses, as noted in §2 above, in arguing that God's existence is necessary if possible. For necessary existence, as Gödel understands it, clearly does correspond to a propositional function of one individual variable. (It is necessary exemplification of the individual's essence(s).[m]) And Gödel's

[m]By relating the necessity thus indirectly to the individual, Gödel avoids quantifying, with an individual variable, into a modal context. Sobel (*1987*, p. 246) cites one exception to the proof's avoidance of this controversial type of quantitification, but the exception is in Scott's version, not in Gödel's **1970*, which uniformly avoids such quantification.

notion of properties is not restricted in its application to any category from which there is an obvious reason for excluding necessary existence.

Of course it does not immediately follow that necessary existence is indeed positive, but there is nothing in Gödel's apparatus to exclude its positiveness. In his **1970* it is asserted as an axiom, but Gödel's notebooks contain at least two arguments for it ("Phil XIV", pages 103–4, 106–7). They are similar to each other; the simpler asserts as axioms that "the necessity of a perfective is a perfective, and being is a perfective" ("Phil XIV", page 106), where "perfective" plays the part played by "positive" in **1970*. From these axioms (fairly plausible on Gödel's assumption that every property, in a broad sense, is either positive or negative), it immediately follows that necessary being is a perfective (positive).

2. Gödel offers several interpretations of the meaning of "positive" (or "perfective"). Only the one that is farthest from his **1970* agrees fully with Leibniz's central idea of the purely positive as involving no negation at all in its construction from simple positive properties. According to the interpretation that seems intended to go with the 1970 proof, "positive means positive in the moral aesthetic sense (independently of the accidental structure of the world)." This classifies "positive" as a value predicate, and indicates that what is positive is necessarily positive,[n] as claimed in Axiom 3 of **1970*. But it does not identify logical properties of positiveness that are likely to be of much help in proving the mutual consistency of all positive properties.

This interpretation also is disturbingly similar to one that is rejected in one of Gödel's notebooks: "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" ("Phil XIV", page 105). The reason given for the rejection, however, is not directed at the assumption that "positive" is a value predicate. The objection is rather that "good" does not express a sufficiently demanding standard of value. That is made clear by the amendment that Gödel goes on to propose: "It is possible to interpret the positive as perfective; that is, 'purely good', that is, such as implies no negation of 'purely good'" ("Phil XIV", page 105). This amendment makes clear that "positive" is to mean *purely* positive or *purely* good, and not just positive or good to some degree.

[n]I take the parenthetical phrase, "independently of the accidental structure of the world", to apply to the positiveness of the positive properties. Perzanowski (*1991*, page 628) seems to take it to apply to any thing's possession of a positive property, for he writes, "According to Gödel, positive means: independent of the accidental structure of the world."

It also specifies an important logical property of (pure) positiveness. Unlike Leibniz, who defined perfections, and purely positive qualities more generally, in terms of the role that negation does not play in their *internal* logical structure, Gödel here characterizes purely positive properties, or "perfectives", in terms of what they *imply*. The importance of this for his ontological proof is underlined as he goes on in his notebook to say, "The chief axiom runs then (essentially): A property is a perfective if and only if it implies no negation of a perfective" ("Phil XIV", page 106). This axiom (or the "only if" half of it) reappears as Axiom 5 in Gödel's **1970*. (The "if" half follows from Axiom 5 together with Axiom 2.) We may reasonably infer that "positive" means *purely* positive in **1970*, and that the "moral aesthetic" explanation of the sense of "positive" given there does not share the feature to which Gödel objected in the rejected explanation in the notebook.

This way of specifying the concept of a (purely) positive property generates the proof of the possibility of God's existence in Gödel's **1970*. Gödel assumes that the sum of all positive properties is itself a positive property (Axiom 1), and that positive properties imply only positive properties (Axiom 5). From these assumptions it follows that "the system of all positive properties is compatible", and hence that the existence of God, as the possessor of all positive properties, is possible.

This possibility proof does not depend on the controversial Leibnizian assumption that the only way in which properties can be incompatible is by formal contradiction arising from negation involved in their construction. That advantage may be outweighed by a major disadvantage, however. If Leibniz's assumptions are accepted, they give a *reason* for believing that all purely positive qualities are mutually consistent, and a sort of explanation of *why* they are consistent, showing that there is no way in which they could be mutually inconsistent. But Gödel's **1970* provides no such explanation, and the axioms from which the mutual compatibility of all purely positive properties is inferred in **1970* are too close to the conclusion to have much probative force to establish it. Of the axiom that "a property is a perfective if and only if it implies no negation of a perfective", Gödel himself, in his notebook, states that it "says essentially that the positive properties form a maximal compatible system" ("Phil XIV", page 106). It seems as fair to say that about Axiom 5 in the 1970 proof. But then is it not question-begging to rely on Axiom 5 to prove that "the system of all positive properties is compatible"?

At the end of **1970* Gödel tersely suggests an alternative, more Leibnizian interpretation of positiveness and a corresponding strategy of proof. Positive, he says, "may also mean pure 'attribution' as opposed to 'privation' (or *containing* privation)". By itself this may be a cryptic formulation, but a footnote explains that what is meant is that "the disjunctive normal form ⟦of a purely positive property⟧ in terms of elementary properties contains a member without negation". Gödel adds that "this interpretation" supports a "simpler proof", but he does not give the proof.

The central idea of the suggested proof is presumably that there is no way in which properties can be mutually inconsistent if the disjunctive normal form of each, in terms of elementary properties, contains at least one member without negation. It must be assumed here that the elementary properties are positive. They correspond to the simple, positive properties of Leibniz's scheme. Gödel sees all other properties as constructed out of them by operations of disjunction (*in*clusive disjunction must be meant here) and negation. Leibniz (if I understand him aright) had allowed no negation at all in the construction of purely positive qualities from simple, positive qualities. Gödel is more liberal on this point, seeing that as long as each purely positive property has in its disjunctive normal form at least one disjunct that involves no negation in its construction, no formal inconsistency can arise among purely positive properties, even if negation is involved in the construction of other disjuncts. In this way he has accomplished an improvement in Leibniz's proof, for the suggested proof seems to have all the advantages of Leibniz's argument, with a less restrictive conception of the purely positive. On the other hand, it depends no less than Leibniz's proof on the controversial assumption that the only way in which properties can be incompatible is by formal contradiction arising from negation involved in their construction.

An even more Leibnizian conception of the purely positive is suggested in Gödel's notebook, when he proposes the theorem: "The positive properties are precisely those that can be formed out of the elementary ones through application of the operations &, $\vee$, $\supset$" ("Phil XIV", page 108). On this construal the purely positive properties will be those that involve no negation at all in their construction from elementary properties (provided the disjunction operation here too is inclusive).

5. Discussion of Gödel's proof, 1970–1991

There is a small but growing secondary literature on Gödel's ontological proof. It has been pointed out that "Gödel's theory is certainly ⟦formally⟧ consistent, having a monistic model comprising one object, one atomic property, hence one ⟦possible⟧ world and, of course, one God."[o] In unpublished work Petr Hájek has proposed proofs of mutual

[o] *Perzanowski 1991*, p. 629.

independence of some of the axioms in Dana Scott's version of the proof.

The first full publication of Gödel's ontological proof was in *Sobel 1987*. Sobel reproduces both Dana Scott's version and Gödel's own **1970*, but discusses chiefly Scott's version. Sobel criticizes the proof as a piece of philosophical theology. One of his main criticisms is that "a being that was *God-like* in the sense of the system would, in connection with many religiously important properties, have not them but their negations." His reason for this claim is that he thinks that some of the traditional attributes of God are incompatible with necessary existence. He deems it "obvious" that no necessarily existing being "would be *sentient* or *cognizant* .... It is at least a firm modal intuition of *mine*", he says, "that there are possible worlds in which there are ... no *sentient* or *cognizant*" things (*Sobel 1987*, pages 249–50).

Sobel's intuitions on this point are shared by many philosophers, but consciously rejected by virtually all partisans of the ontological argument. It would be naive to expect the latter to accept Sobel's objection and conclude that God is not a cognizant being. The form of Sobel's objection is therefore somewhat misleading. Friends of the ontological argument are bound to see it as merely a repackaging of a familiar empiricist objection, based on the claim (consciously rejected by them) that a being possessing the sort of reality generally ascribed to God could not exist necessarily.

It remains a serious question, however, whether the being whose existence is purportedly proved by Gödel's ontological proof is the God of traditional theism. Despite its role in the philosophical theologies of Leibniz, Wolff, and Kant, and its resonance with many medieval philosophical theologies, it is not immediately obvious that the concept of a being possessing the sum of all purely positive properties (or qualities) is a concept of God. Any employment of Gödel's ontological proof in philosophical theology would require further argument on this point, with particular attention to Gödel's conception of positive properties.

Sobel's other main objection is that the assumptions of Gödel's ontological proof generate a proof that all truths are necessary truths. For on a liberal construal of the notion of a property, "if something is true, then ... a God-like being ⟦if one actually exists⟧ has *the property of being in the presence of this truth.* But every property of a God-like being ⟦i.e., every actual property of God⟧ is necessarily instantiated, from which it follows that this truth ⟦i.e., any actual truth⟧ is a necessary truth."[p] (That every actual property of God is necessarily instantiated follows from $N(\exists y)G(y)$, the conclusion of Gödel's ontological proof, since G,

[p] *Sobel 1987*, p. 253. Sobel also gives a formal proof based on the ideas contained in this informal exposition.

as the "essence" of God, in Gödel's sense, entails all of God's actual properties.)

It is characteristic of Leibnizian philosophical theology to be in some danger of leaving no truths contingent (see *Adams 1977*). And it is not altogether clear that Gödel was determined to avoid such a necessitarian conclusion. He wrote a notebook entry about the ontological proof in which he seems quite favorable to the thesis that "for every compatible system of properties there is a thing" ("Phil XIV", page 107). That thesis looks strongly necessitarian, but the interpretation of the entry containing it is not obvious; one may wonder, for instance, whether merely possible objects count as "things" here.[q]

Another relevant entry comes at the very end of the notebook section devoted to the ontological proof. Gödel had written that propositions "of the form $\varphi(a)$" are "the only synthetic propositions" because "they depend not on God, but on the thing a" ("Phil XIV", page 108). In this context a must be an individual other than God. Such individuals, and truths about them, do depend causally on God, according to traditional theism. The independence Gödel has in mind here is presumably logical rather than causal. Later, at the bottom of the following page, with a line indicating insertion at this point, or reference to it, Gödel wrote ("Phil XIV", page 109):

> This doesn't work, because then God would have an imperfective, which consists in the fact that imperfectives are possible. Everything that follows from a perfective, such as something good, that is a perfective, is.

This correction, I think, must have arisen from something like the following train of thought: If my having gray hair is synthetic and logically independent of God, then it is contingent, and both it and its falsity are possible, and likewise for your having whatever color of hair you have. Gödel seems to take these possibilities as implying "that imperfectives are possible"—presumably on the ground that of the two properties, (1) having gray hair and (2) not having gray hair, one must be a perfective and the other an imperfective.

But why would "the fact that imperfectives are possible" constitute an imperfective that God would have? Here I suppose we must invoke something like Sobel's assumption that, for every truth, God has the property of coexisting with that truth, or perhaps the traditional theistic assumption that, for every truth, God has the property of knowing that truth. (These assumptions imply that my having gray hair is not, after all, logically independent of all of God's properties; perhaps the independence Gödel had in mind is only a logical independence from God's internal, nonrelational properties; or perhaps it is an independence from God's necessary properties, assuming for the sake of the present argument that God may have some contingent properties.) Then since I in fact have gray hair, God has in fact the property of coexisting with my having gray hair. And that property must be a perfective if all God's (actual) properties are perfectives. Its negation, the property of not coexisting with my having gray hair, must then be an imperfective, given Gödel's assumption that every negation of a perfective is an imperfective. But if my having gray hair is possibly false, then God has the property of possibly not coexisting with my having gray hair. And this possibility will be an imperfective; for, as Gödel maintains in his notebook ("Phil XIV", page 103*n*; cf. page 107), the possibility of a negative is negative, and presumably the possibility of an imperfective must also be imperfective. So if God has no imperfectives, as Gödel's definition of deity requires, my having gray hair (when I do) must not be possibly false, and in general, "everything that follows from a perfective, such as something good", must be—a conclusion of Leibnizian optimism, and perhaps more than Leibnizian necessitarianism.

[q]As Charles Parsons has suggested to me might be the case.

I grant that the suggestion that Gödel would have accepted the sweeping necessitarian implication with which Sobel charges him is somewhat speculative. In any event, there are possible modifications of Gödel's assumptions that avoid the sweeping necessitarianism without undermining his ontological proof. Axiom 2 of his **1970* is equivalent to the conjunction of two conditionals:

(A) If a property is positive, then its negation is not positive.

(B) If a property is not positive, then its negation is positive.

Anderson (*1990*) has pointed out that of these conditionals, only (A) is required for Gödel's ontological proof, but (B) is required for the proof that all truths are necessary. He argues that (B) is less plausible than (A), as (B) "seems to overlook a possibility: that both a property and its negation should be *indifferent*". He sets out a revised version of Gödel's ontological proof, which has (A) but not (B) as an axiom, and which still has the conclusion that "the property of being God-like* is necessarily exemplified". Anderson's version of the proof also differs from Gödel's in not requiring an essence of a thing to entail all the actual properties of the thing, but only a subset classified as "essential" to the thing, and in defining a God-like* being as one that has all and only the positive properties as *essential* properties, and not merely as properties. Anderson proposes a "possible worlds" model to prove that the assumptions of this proof are consistent with there being contingent truths (*Anderson 1990*, pages 295–97).

Another way of avoiding the sweeping necessitarian conclusion is to use a more restrictive notion of a property than Sobel does. In deriving the necessitarian conclusion, he relies on the very strong assumption that "properties" include all those abstracted in accordance with the principle

$$\hat{\beta}[\varphi](\alpha) \equiv \varphi',$$

"where β is an individual variable, α is a term, φ is a formula, and φ' is a formula that comes from φ by proper substitution of α for β" (*Sobel 1987*, page 251). This assumption is not part of Gödel's argument, and Hájek, in the unpublished work cited above, has argued that if it is replaced with certain weaker assumptions about properties, the axioms of (Scott's version of) Gödel's ontological proof can be shown by a "possible worlds" model to be consistent with the existence of contingent truths.

One way of filling out Hájek's suggestion would be to restrict the category of "properties" to *nonrelational* properties for purposes of the ontological proof. This could be accomplished by restricting it to properties that Leibniz would have counted as *qualities*, but it might not be necessary to go that far. The important point is that if relational properties are not counted as properties for purposes of the argument, then such "properties" as that of "being in the presence of this truth", which are relational, will not be among the actual properties of God that must be necessarily instantiated according to the argument. If the properties of God that are necessarily instantiated are exclusively nonrelational, then their necessity will not imply the necessity of truths about other beings.

Robert Merrihew Adams[r]

[r]I am indebted to Charles Parsons for helpful comments on an earlier version of this note, and to Jay Atlas, Dana Scott, and Morton White for sharing their recollections bearing on the history of Gödel's ontological proof.

Ontological proof
(*1970)

Feb. 10, 1970

$P(\varphi)$ φ is positive (or $\varphi \in P$).

Axiom 1. $P(\varphi).P(\psi) \supset P(\varphi.\psi)$.[1]

Axiom 2. $P(\varphi) \vee P(\sim\varphi)$.[2]

Definition 1. $G(x) \equiv (\varphi)[P(\varphi) \supset \varphi(x)]$ (God)

Definition 2. $\varphi\,\mathrm{Ess}.\,x \equiv (\psi)[\psi(x) \supset N(y)[\varphi(y) \supset \psi(y)]]$. (Essence of x)[3]

$$p \supset_N q \;=\; N(p \supset q). \quad \text{Necessity}$$

Axiom 3.
$$P(\varphi) \supset NP(\varphi)$$
$$\sim P(\varphi) \supset N\sim P(\varphi)$$

because it follows from the nature of the property.[a]

Theorem. $G(x) \supset G\,\mathrm{Ess}.x$.

Definition. $E(x) \equiv (\varphi)[\varphi\,\mathrm{Ess}\,x \supset N(\exists x)\,\varphi(x)]$. (necessary Existence)

Axiom 4. $P(E)$.

Theorem. $G(x) \supset N(\exists y)G(y)$,
hence $(\exists x)G(x) \supset N(\exists y)G(y)$;
hence $M(\exists x)G(x) \supset MN(\exists y)G(y)$. ($M$ = possibility)
$M(\exists x)G(x) \supset N(\exists y)G(y)$.

$M(\exists x)G(x)$ means the system of all positive properties is compatible. 2
This is true because of:
Axiom 5. $P(\varphi).\varphi \supset_N \psi :\supset P(\psi)$, which implies

$$\begin{cases} x = x & \text{is positive} \\ x \neq x & \text{is negative.} \end{cases}$$

[1]And for any number of summands.

[2]Exclusive or.

[3]Any two essences of x are *necessarily equivalent*.

[a]Gödel numbered two different axioms with the numeral "2". This double numbering was maintained in the printed version found in *Sobel 1987*. We have renumbered here in order to simplify reference to the axioms.

But if a system S of positive properties were incompatible, it would mean that the sum property s (which is positive) would be $x \neq x$.

Positive means positive in the moral aesthetic sense (independently of the accidental structure of the world). Only then ⟦are⟧ the axioms true. It may also mean pure "attribution"[4] as opposed to "privation" (or *containing* privation). This interpretation ⟦supports a⟧ simpler proof.

If φ ⟦is⟧ positive then *not*: $(x)N{\sim}\varphi(x)$. Otherwise: $\varphi(x) \supset_N x \neq x$; hence $x \neq x$ ⟦is⟧ positive, so $x = x$ ⟦is⟧ negative, contrary ⟦to⟧ Axiom 5 or the existence of positive properties.

---

[4]I.e., the disjunctive normal form in terms of elementary properties[b] contains a member without negation.

[b]Here Gödel uses the abbreviation "prop.", which could be read, in isolation, either as "properties" or "propositions". In the context, however, it is clear that it is properties whose positiveness is under discussion. The related discussion in the excerpts from "Phil XIV" in the appendix, below, explicitly concerns "positive properties". With regard to fn. 4, where the reference to "disjunctive normal form" might lead us to think first of propositions, note that in "Phil XIV", p. 108, Gödel speaks explicitly of properties ("Eigenschaften") that are "members of the conjunctive normal form" of complex properties. An interpretation of fn. 4 is offered in the introductory note, pp. 397–398 above.

Introductory note to *Gödel *1970a*, **1970b*, and **1970c*

1. Introduction

*Gödel *1970a* is a handwritten document that was sent to Alfred Tarski for submission to the *Proceedings of the National Academy of Sciences*. It lists four axioms and claims to deduce from them that $2^{\aleph_0} = \aleph_2$. **1970b* and **1970c* are two other handwritten documents that bear on **1970a*. **1970b* claims, on the contrary, to deduce the continuum hypothesis from some of the axioms mentioned in **1970a*. **1970c* is a letter to Tarski (apparently never sent; cf. Gregory Moore's introductory note to *Gödel 1947* and *1964*, these *Works*, Volume II, page 175) which acknowledges serious errors in **1970a*.[a]

Upon receiving **1970a*, Tarski asked the author to referee the manuscript. After a careful study of the manuscript, I was unable to follow the argument. I reported back to Tarski that if the author were anyone but Gödel, I would certainly recommend that the manuscript be rejected. (The manuscript has never been published hitherto.) Subsequently, D. A. Martin showed that a key argument of the paper was demonstrably wrong. (See section 6 below.)[b]

In the remainder of this introductory note, I shall describe the proof in **1970a* to the extent that I now understand it. I shall also discuss a model of set theory whose study sheds considerable light on **1970a*. What comments I make on the other papers will be in the course of discussing **1970a*.

The rest of this note is organized as follows: Section 2 gives my reconstruction of the precise formulation of the four axioms from which **1970a* purports to deduce that $2^{\aleph_0} = \aleph_2$. Section 3 contains an outline of Gödel's alleged proof. Section 4 contains the proof that Axioms 3 and 4 entail $2^{\aleph_0} = 2^{\aleph_1}$. Section 5 presents the main result of **1970b*: The rectangular axiom $A(\aleph_1, \aleph_0)$ implies $2^{\aleph_0} = \aleph_1$. Section 6 is devoted

[a]The three papers bear (in Gödel's handwriting) the notations "I Fassung", "II Fassung", and "III Fassung". ("Fassung" is German for "version".) Thus it seems likely that Gödel viewed them as different installments or versions of a single paper. In addition, **1970b* has written at its top "Nur für mich geschrieben" ("Written only for me").

[b]For further discussion of the history of **1970a*, we recommend Moore's introductory note to *Gödel 1947* and *1964*, especially pp. 173–175.

The paper *Ellentuck 1975* discusses many of the issues discussed in the present note, and the interested reader might wish to read it. In addition, *Takeuti 1978* explores some interesting generalizations of the arguments of section 5.