A Study of Gödel's Ontological Argument

Andrew Jackson [1,*]
[1] Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom

I. INTRODUCTION

Kurt Gödel is considered to be one of the greatest logicians of all time. He is best known for two important developments in the history of logic: his completeness theorem [1] (proven in 1929 as part of his doctoral thesis [2]) and his incompleteness theorems [3, 4] (proven in 1931, which put an end to both Hilbert's program [5] and Whitehead and Russell's project in Principia Mathematica [6]) - due to which which John von Neumann claimed "the subject of logic has certainly completely changed its nature and possibilities" [7].

His work in mainstream logic was "monumental" [7] but perhaps his most ambitious work was his ontological argument; a claimed proof of the existence of God (defined in a slightly unorthodox way).

This claimed proof is believed to have been written - in its earliest form - in the early 1940s, however it only became known to a wider audience (than just Gödel himself) in the 1970s. However it was not published until 1987. Gödel himself appears to have accepted the argument and considered himself "theistic, not pantheistic, following Leibniz rather than Spinoza" [8]: a viewpoint consistent with the God claimed by his ontological argument.

Herein I present and examine Gödel's Ontological Argument in an atypical order[1], to allow for closer examination of the axioms and hopefully a more pedagogical presentation of the definitions and argument. Furthermore, I provide natural deductions of all theorems.
I assume only a knowledge of first-order logic [9] and its associated Gentzen-style natural deduction system [10, 11].

There are several known problems with the axioms of Gödel's ontological argument. Such as that the axioms are known to imply modal collapse [12] (i.e. that every true statement becomes necessarily true and there are no contingent truths), among others.

II. CHOOSING A LOGIC

Definition 1. A modal formula is any formula that contains either $\Diamond$ or $\Box$. Any formula that does not contain either is modal-free.

The first issue in working formally with any argument is the choice of which logic to use to tackle it. As Gödel's ontological argument is meant to be an argument about what could possibly be the case, it makes sense to use a modal logic where:

1. All tautologies are true,
2. Modus ponens holds. I.e. if A and $A \to B$, then B.

This motivates the use of a modal logic.

Definition 2. A modal logic is a set, Σ, of modal formulas such that:

1. Σ contains all tautologies,
2. If $A \in \Sigma$ and $D_1, D_2, ..., D_n$ are formulas, then:
$$A[D_1/p_1, D_2/p_2, ..., D_n/p_n] \in \Sigma, \tag{1}$$
3. If $A \in \Sigma$ and $A \to B \in \Sigma$, then $B \in \Sigma$.

I then look for other properties that may be required/convenient for the derivation. Examples of these are:

1. The standard relationship between possibility and necessity. I.e. that something being possible is equivalent to it not being neccessary that it is not true. This can be expressed formally as:
$$\Diamond p \leftrightarrow \neg\Box\neg p. \tag{2}$$
2. Anything that can be derived logically, purely from the logic, is necessarily true.
3. All derivations are necessarily true.

This motivates the use of a normal modal logic.

Definition 3. A normal modal logic, Σ, is a modal logic such that: $\forall p, q$,

1. $\Box(p \to q) \to (\Box p \to \Box q)$,
2. $\Diamond p \leftrightarrow \neg\Box\neg p$,
3. $A \in \Sigma$ implies $\Box A \in \Sigma$.

The smallest normal modal logic - i.e. the minimal set of modal formulas satisfying the conditions to be a normal modal logic - is denoted as **K**. I can add formulas to **K** to construct larger modal logics. In fact, we have good reason to do so, but before looking at what we should add to **K**, we consider the approach and formalism to use.

Definition 4. A model for a modal language is a triple: $\mathcal{M} = \langle W, R, V\rangle$, where:

1. W is a non-empty set of "worlds",

* Andrew.J.Jackson@warwick.ac.uk
[1] There is a standard presentation that mixes the axioms, definitions, and theorems but here I separate them

2. R is a binary accessibility relation on W,
3. V is a function assigning to each proposition, p, a set of worlds, $V(p)$.

I then define how we model truth.

Definition 5. A formula, A, is said to be true in a world, ω, in model, $\mathcal{M}$, - denoted $\mathcal{M}, \omega \vDash A$ - if it satisfies the following inductive definition:

1. For all possible worlds, ω, in all possible models, $\mathcal{M}$, $\mathcal{M}, \omega \nvDash \bot$.
2. $\mathcal{M}, \omega \vDash p \iff \omega \in V(p)$.
3. $\mathcal{M}, \omega \vDash \neg p \iff \mathcal{M}, \omega \nvDash p$.
4. $\mathcal{M}, \omega \vDash (p \wedge q) \iff \mathcal{M}, \omega. \vDash p$ and $\mathcal{M}, \omega \vdash_\Sigma q$.
5. $\mathcal{M}, \omega \vDash (p \vee q) \iff \mathcal{M}, \omega \vDash p$ or $\mathcal{M}, \omega \vDash q$.
6. $\mathcal{M}, \omega \vDash (p \rightarrow q) \iff \mathcal{M}, \omega \nvDash p$ or $\mathcal{M}, \omega \vdash_\Sigma q$.
7. $\mathcal{M}, \omega \vDash \Box p \iff \mathcal{M}, \omega' \vDash p$ for all $\omega' \in W$ such that $R\omega\omega'$.
8. $\mathcal{M}, \omega \vDash \Diamond p \iff \mathcal{M}, \omega' \vDash p$ for at least one $\omega' \in W$ such that $R\omega\omega'$.

From this, it becomes clear that the point of the R binary relation is to define which worlds are "accessible" from each other. Given the nature of my claim - and reading Gödel's original work - it seems sensible to adopt a model with universal accessibility: i.e. $\forall \omega, \omega' \in W, R\omega\omega'$. This then points to a particular normal modal logic we should use: the minimal normal modal logic corresponding to a model with universal accessibility.

III. SUBJECT-MATTER DEFINITIONS

A. Positive Properties

B. Definition 1: Defining God and God-like Objects

1. Logical Expression of Definition One

Definition one is denoted as **Df.1** and can be expressed logically as:

$$G(x) \leftrightarrow \forall \phi(P(\phi) \rightarrow \phi(x)). \tag{3}$$

2. Informal Expression and Discussion of Definition One

The purpose of definition one is to define God-like objects. An object x is denoted as being God-like by the expression $G(x)$ and this is defined, in Eqn. 3, as object x having all positive properties.

The whole purpose of Gödel's ontological argument is to attempt to establish the existence of some object that meets this definition. It therefore follows that the correctness of this definition is essential.

How well definition one serves as a definition of God is a very deep question. One that I'm not nearly even close to having the slightest chance in hell of answering. Talk to a priest/rabi/imam of your choice for more details.

C. Definition 2: Defining Essence

1. Logical Expression of Definition Two

Definition two is denoted as **Df.2** and can be expressed logically as:

$$\phi \; ess \; x \leftrightarrow \phi(x) \wedge \forall \psi(\psi(x) \rightarrow \Box \forall y(\phi(y) \rightarrow \psi(y))). \tag{4}$$

2. Informal Expression and Discussion of Definition Two

This notion of an object's essence breaks into two parts:

1. $\phi(x)$
2. $\forall \psi(\psi(x) \rightarrow \Box \forall y(\phi(y) \rightarrow \psi(y)))$

The first ($\phi(x)$) that, for ϕ to be the essence of x, x must have the property ϕ, I do not think requires much arguing; it could hardly not be the case. The second part requires more consideration. It states that the essence of an object is *the* defining feature of that object and implies all other properties of *any* other object that has that property (the essence). In fact this definition implies that the essence of an object uniquely and completely defines it.

The two parts of the definition may be combined to read: an essence of an object is a property possessed by it and implying any and all of its properties.

Objects can be uniquely defined by listing all their properties and only their properties.

I note that the ontological argument does not require the reader accept that this definition captures what an essence means (to them), or that it conforms to any other pre-existing conception they may have. The reader just needs to accept that we are using $\phi \; ess \; x$ as a shorthand for $\phi(x) \wedge \forall \psi(\psi(x) \rightarrow \Box \forall y(\phi(y) \rightarrow \psi(y)))$.

D. Definition 3: Defining Necessary Existence

1. Logical Expression of Definition Three

Definition three is denoted as **Df.3** and can be expressed logically as:

$$E(x) \leftrightarrow \forall \phi(\phi \; ess \; x \rightarrow \Box \exists y \phi(y)). \tag{5}$$

2. Informal Expression and Discussion of Definition Three

Definition three defines necessary existence, denoted as E. An object necessarily exists if and only if for any property that is an essence of that object, in all possible worlds, an object with that property exists.

The key question around necessary existence is can any object with this property exist? As for any property an object it is the essence of can be constructed, then for any and all properties that imply necessary existence there must exist an object with that property as its essence in every possible world. However, necessary existence as a concept originates in Gödel's ontological argument: it was not and has not been found in nature. There are three ways to resolve this incongruence:

1. necessary existence is a contradiction and so objects with necessary existence cannot exist
2. properties that imply necessary existence are rare, or non-existent, to the point where such objects have not been found
3. Some consequence of necessary existence hides objects with it from us finding them

I note that all but the first of these conclusions is not a problem but it does raise the possibility that there may be issues with the concept of necessary existence.

Again I note that the ontological argument does not require the reader accept that this definition conforms to any other pre-existing conception they may have. The reader just needs to accept that we are using $E(x)$ as a shorthand for $\forall\phi(\phi\ ess\ x \rightarrow \Box\exists y\phi(y))$.

E. Denoting the Set of All Definitions

I denote the set of all the definitions by **Dfs**.

IV. AXIOMS

A. Axiom One

1. Logical Expression of Axiom One

The logical expression of axiom one is denoted as **Ax.1** and is:

$$(P(\phi) \land \Box\forall x(\phi(x) \rightarrow \psi(x))) \rightarrow P(\psi). \tag{6}$$

2. Informal Expression and Discussion of Axiom One

Axiom one can be informally stated as any property implied by a positive property in all possible world (i.e. is necessarily implied by it) is a positive property.

This is - I believe - not an easy axiom to justify. Some comments can be made to support it though. If a positive property necessarily implies another property, then whenever the positive property is present so is the other - implied - property. Therefore, the second (implied) property can be see as a feature of the positive property (that implies it) and so can impact the positivity of the positive property i.e. if the implied property were so bad that it completely overwhelmed the gains or the positives from having the positive property then - by virtue of implying the overwhelming negativity of the necessarily implied property - the positive property would not be positive. As the positive property is positive, this is a contradiction; so I conclude that being implied necessarily by a positive property at least lower bounds how negative a property can be. This is far from proving that the implied property is positive but it does weakly support Axiom one.

On the other hand, there may be counter-examples to Axiom one, and without straying too far from our focus on God. Often, whenever the question of why God allows bad things to happen, the response is that it is the cost of giving humans free will. If you accept this argument and that there is a clear inference that free will is a positive property of people, and that free will necessarily implies the possibility that humans will do bad things, then - as it seems hard to argue that the ability to do immoral/harmful things is a positive property of humans - this serves as a counter-example to Axiom one.

However, if we take a stricter notion of what is a positive property, such as insisting that a property is only positive if it does not necessarily imply any property that is not positive, then Axiom one seems a lot more reasonable. However this does make it harder to argue that a given property is positive.

B. Axiom Two

1. Logical Expression of Axiom Two

The logical expression of axiom two is denoted as **Ax.2** and is:

$$P(\neg\phi) \leftrightarrow \neg P(\phi). \tag{7}$$

2. Informal Expression and Discussion of Axiom Two

Eqn. 7 may be re-expressed - perhaps more easily examinably - as:

$$[\neg(P(\phi) \land P(\neg\phi))] \land [P(\neg\phi) \lor P(\phi)]. \tag{8}$$

Eqn. 8 - and hence Axiom two - may be expressed informally as: any property is either positive or it's negation is, but not both.

Eqn. 8 also allows me to break Axiom two into two parts and defend each separately (although both are required for Axiom 2 to hold). The first part is: $\neg(P(\phi) \land P(\neg\phi))$, that for any property, ϕ, I cannot accept both ϕ and $\neg\phi$ being positive. To examine this, assume ϕ is positive and consider that if it is good to have a property (i.e. it is positive), then not having it is worse than having it, hence - in isolation - not having that property is not a positive property. So $\neg\phi$ cannot be positive.

If I assume $\neg\phi$ is positive then the same argument implies that ϕ cannot be positive.

The second part of Axiom two that needs considering is $P(\neg\phi) \vee P(\phi)$ i.e. that at least one of ϕ and $\neg\phi$ must be positive. This is harder to defend. A naive defence may be that I must have some preference between ϕ and $\neg\phi$ and that preference entails that one is positive but this runs in the vagueness of how positivity is defined in the ontological argument[2]: does your preference translate into postivity in any way? I also take issue that I may have a preference between *any* two properties, even if I have all details of the property ϕ, all implications of it, and the entire universe. ϕ may be some utterly irrelevant property and so no preference can be chosen.

C. Axiom Three

1. Logical Expression of Axiom Three

The logical expression of axiom three is denoted as **Ax.3** and is: $P(G)$.

2. Informal Expression and Discussion of Axiom Three

Axiom three is fairly simple, it simply states that being God-like is a positive property. Axiom three is - I believe - fairly easy to defend: being God-like is defined as having all positive qualities and so it stands to reason it is itself positive. It would be strange if having *all* positive qualities was not positive.

I also note that if I accept Axiom two and that having all positive properties is not positive, then not having all positive properties *is* positive. This then implies that any object that has all positive properties has the property that it does not have all positive properties. I then have to conclude that either no object can have all properties (in which case, there is no point to Gödel's ontological argument as God is already precluded from existing) or that having all positive properties is positive.

D. Axiom Four

1. Logical Expression of Axiom Four

The logical expression of axiom four is denoted as **Ax.4** and is:

$$P(\phi) \rightarrow \Box P(\phi). \tag{9}$$

2. Informal Expression and Discussion of Axiom Four

Axiom Four may be expressed informally as: any property that is positive is necessarily positive. I.e. a positive property is positive in all possible worlds. I think Axiom four is another relatively easy one to defend: it essentially claims that positive properties - e.g. being moral - are not positive by chance; there was not a contingent event that made them positive.

This axiom may be contested on consequentialist grounds as what if there are contingent circumstances that could have rendered the consequences or implications of a given positive property inconsequential or negative? E.g. an object may have some amazingly positive property but its positive effects are going to be non-existence if it is destroyed immediately after coming into existence.

However, this may be countered by arguing that *only* properties that could not possibly - in any possible world - not be positive are positive. I.e. raising the bar of positivity. An example of a property meeting this new bar is the property of having all positive properties: there is no contingent situation where this could be negative (or neutral), and Axiom two would imply this makes it positive necessarily.

E. Axiom Five

1. Logical Expression of Axiom Five

The logical expression of axiom five is denoted as **Ax.5** and is: $P(E)$.

2. Informal Expression and Discussion of Axiom Five

An informal expression of Axiom five is that necessary existence is a positive property i.e. it is positive to have any property that constitutes your essence necessarily imply an object exists with that property. However there is no guarantee that that object is you; hence necessary existence does not guarantee that you exist, just that an object with your essence does. In which case, your necessary existence does not positively affect you - if there even is a you.

However there is an object with that property, so if your essence is positive then your necessary existence implies an object with a positive property necessarily exists, so this could be good in many cases.

F. Denoting the Set of All Axioms

I denote the set of all the definitions by **Axs**. Likewise, I denote the set of all theorems up to and including the nth (where $n \in \mathbb{N}$) by $\mathbf{Ths}_n$.

[2] An issue I am not inclined to fix as ....

V. NATURAL DEDUCTION OF GÖDEL'S ONTOLOGICAL ARGUMENT

I would like to note that I use the phrase reasoning within a world (rww) to denote applying logical derivations of B from A within a noted set (possibly all or just one) of worlds where A is true to imply that B is true within all those worlds too.

A. Derivation of Theorem 1

Informal Statement of Theorem 1. For any positive property, there is at least one possible world where there exists an object with that property.

Theorem 1. $\boldsymbol{Dfs}, \boldsymbol{Axs} \vdash_\Sigma P(\phi) \to \Diamond\exists x\phi(x)$

Proof. Let δ' be the following derivation:

$$\scriptstyle(1)\ \dfrac{\dfrac{\dfrac{\forall x\neg\phi(x)}{\neg\phi(a)}\ \forall\text{ Elim} \qquad \dfrac{[\exists x\phi(x)]^1}{\phi(a)}\ \exists\text{ Elim}}{\bot}\ \neg\text{ Elim}}{\neg\exists x\phi(x)}\ \bot_c$$

This derivation shows that $\forall x\neg\phi(x) \vdash_\Sigma \neg\exists x\phi(x)$ and is then used within a world in the below derivation showing the full theorem.

$$(3)\ \dfrac{\delta\ \dfrac{\dfrac{(2)\ \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{(1)\ \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\Box\forall x\neg\phi(x)]^2}{\forall x\neg\phi(x)}\ \Box\text{ Elim}}{\neg\phi(a)}\ \forall\text{ Elim} \qquad [\phi(a)]^1}{\bot}\ \neg\text{ Elim}}{\neg\phi(a)}\ \bot_I}{\phi(a) \to \neg\phi(a)}\ \to\text{ Intro}}{\forall x(\phi(x) \to \neg\phi(x))}\ \forall\text{ Intro} \qquad [P(\phi)]^3}{P(\phi) \wedge \forall x(\phi(x) \to \neg\phi(x))}\ \wedge\text{ Intro} \qquad \textbf{Ax.1}}{P(\neg\phi)}\ \to\text{ Elim} \qquad \textbf{Ax.2}}{\neg P(\phi)}\ \to\text{ Elim} \qquad [P(\phi)]^3}{\bot}\ \neg\text{ Elim}}{\neg\Box\forall x\neg\phi(x)}\ \bot_c}{\neg\Box\neg\exists x\phi(x)}\ \text{rww, } \Diamond \equiv \neg\Box\neg}{\Diamond\exists x\phi(x)}\ \text{Using condition 2 of Def. 3}}{P(\phi) \to \Diamond\exists x\phi(x)}\ \to\text{ Intro}$$

□

B. Derivation of Theorem 2

Informal Statement of Theorem 2. There exists a possible world where a God-like object exists.

Theorem 2. $\boldsymbol{Dfs}, \boldsymbol{Axs}, \boldsymbol{Ths}_1 \vdash_\Sigma \Diamond\exists x G(x)$.

Proof. This proof is a fairly straightforward application of

$$\dfrac{\textbf{Th.1} \qquad \textbf{Ax.3}}{\Diamond\exists x G(x)}\ \to\text{ Intro}$$

□

C. Derivation of Theorem 3

Informal Statement of Theorem 3. For any God-like object, being God-like is an essential property of that object.

Theorem 3. $\boldsymbol{Dfs}, \boldsymbol{Axs} \vdash_{\Sigma} G(x) \to G\ ess\ x.$

Proof. Let δ_2 denote the natural deduction:

$$
\begin{prooftree}
\AxiomC{$\textbf{Df.1}$}
\AxiomC{$[G(a)]^1$}
\RightLabel{$\to$ Elim}
\BinaryInfC{$\forall\phi(P(\phi) \to \phi(a))$}
\RightLabel{$\forall$ Elim}
\UnaryInfC{$P(\Gamma) \to \Gamma(a)$}
\AxiomC{$P(\Gamma)$}
\RightLabel{$\to$ Elim}
\BinaryInfC{$\Gamma(a)$}
\LeftLabel{(1)}
\RightLabel{$\to$ Intro}
\UnaryInfC{$G(a) \to \Gamma(a)$}
\RightLabel{$\forall$ Intro}
\UnaryInfC{$\forall y(G(y) \to \Gamma(y))$}
\end{prooftree}
$$

I'll skip the natural deductions (as they are standard textbook derivations) but let δ_2' be the deduction that for any property, Γ, $\neg\neg\Gamma = \Gamma$ and δ_2'' be that an implication implies its contra-positive.

$$
\begin{prooftree}
\AxiomC{$\textbf{Df.1}$}
\AxiomC{$[G(x)]^1$}
\BinaryInfC{$\forall\phi(P(\phi) \to \phi(x))$}
\RightLabel{$\forall$ Elim}
\UnaryInfC{$P(\neg\Gamma) \to \neg\Gamma(x)$}
\RightLabel{δ_2''}
\UnaryInfC{$\Gamma(x) \to \neg P(\neg\Gamma)$}
\AxiomC{$[\Gamma(x)]^2$}
\RightLabel{$\to$ Elim}
\BinaryInfC{$\neg P(\neg\Gamma)$}
\AxiomC{$\textbf{Ax.2}$}
\BinaryInfC{$\neg\neg P(\Gamma)$}
\RightLabel{δ_2'}
\UnaryInfC{$P(\Gamma)$}
\AxiomC{$\textbf{Ax.4}$}
\BinaryInfC{$\Box P(\Gamma)$}
\RightLabel{rww, δ_2}
\UnaryInfC{$\Box\forall y(G(y) \to \Gamma(y))$}
\LeftLabel{(2)}
\RightLabel{$\to$ Intro}
\UnaryInfC{$\Gamma(x) \to (\Box\forall y(G(y) \to \Gamma(y)))$}
\RightLabel{$\forall$ Intro}
\UnaryInfC{$\forall\psi(\psi(x) \to (\Box\forall y(G(y) \to \psi(y))))$}
\AxiomC{$[G(x)]^1$}
\RightLabel{$\wedge$ Intro}
\BinaryInfC{$G(x) \wedge \forall\psi(\psi(x) \to (\Box\forall y(G(y) \to \psi(y))))$}
\AxiomC{$\textbf{Df.2}$}
\BinaryInfC{$G\ ess\ x$}
\LeftLabel{(1)}
\RightLabel{$\to$ Intro}
\UnaryInfC{$G(x) \to G\ ess\ x$}
\end{prooftree}
$$

$\square$

D. Derivation of Theorem 4

Informal Statement of Theorem 4. There exists a God-like object in every possible world.

Theorem 4. $\boldsymbol{Dfs}, \boldsymbol{Axs}, \boldsymbol{Ths}_3 \vdash_{\Sigma} \Box\exists y G(y).$

Proof.

$$
\begin{prooftree}
\AxiomC{$[G(a)]^1$}
\AxiomC{$\textbf{Df.1}$}
\RightLabel{$\to$ Elim}
\BinaryInfC{$\forall\phi(P(\phi) \to \phi(a))$}
\RightLabel{$\forall$ Elim}
\UnaryInfC{$P(E) \to E(a)$}
\AxiomC{$\textbf{Ax.5}$}
\RightLabel{$\to$ Elim}
\BinaryInfC{$E(a)$}
\AxiomC{$\textbf{Th.3}$}
\AxiomC{$[G(a)]^1$}
\RightLabel{$\to$ Elim}
\BinaryInfC{G ess a}
\RightLabel{$\to$ Elim}
\BinaryInfC{$\Box\exists y G(y)$}
\LeftLabel{(1)}
\RightLabel{$\to$ Intro}
\UnaryInfC{$G(a) \to \Box\exists y G(y)$}
\RightLabel{$\exists$ Intro}
\UnaryInfC{$\exists x G(x) \to \Box\exists y G(y)$}
\RightLabel{Lemma 1}
\UnaryInfC{$\neg\Box\exists y G(y) \to \neg\exists x G(x)$}
\RightLabel{All derivations are necessary}
\UnaryInfC{$\Box(\neg\Box\exists y G(y) \to \neg\exists x G(x))$}
\RightLabel{Condition 1 of normal modal logics}
\UnaryInfC{$(\Box\neg\Box\exists y G(y)) \to (\Box\neg\exists x G(x))$}
\RightLabel{Lemma 1}
\UnaryInfC{$(\neg\Box\neg\exists x G(x)) \to (\neg\Box\neg\Box\exists y G(y))$}
\RightLabel{$\neg\Box\neg \equiv \Diamond$}
\UnaryInfC{$\Diamond\exists x G(x) \to \Diamond\Box\exists y G(y)$}
\AxiomC{$\textbf{Th.2}$}
\RightLabel{$\to$ Elim}
\BinaryInfC{$\Diamond\Box\exists y G(y)$}
\RightLabel{Lemma 2}
\UnaryInfC{$\Box\exists y G(y)$}
\end{prooftree}
$$

□

VI. DISCUSSION

VII. ACKNOWLEDGEMENTS

[1] K. Gödel, Die vollständigkeit der axiome des logischen funktionenkalküls, Monatshefte für Mathematik und Physik **37**, 349 (1930).

[2] K. Gödel, gerÜber die vollständigkeit des logikkalküls (1929).

[3] K. Gödel, Über formal unentscheidbare sätze der principia mathematica und verwandter systeme i, Monatshefte für Mathematik und Physik **38**, 173 (1931).

[4] S. Batzoglou, Goedel's incompleteness theorem (2021), arXiv:2112.06641 [math.HO].

[5] R. Zach, Hilbert's program then and now, in *Philosophy of Logic* (Elsevier, 2007) p. 411–447.

[6] A. N. Whitehead and B. Russell, *Principia Mathematica* (Cambridge University Press, Cambridge, 1910–1913).

[7] P. R. Halmos, The legend of john von neumann, The American Mathematical Monthly **80**, 382 (1973).

[8] H. Wang, *Reflections on Kurt Gödel* (MIT Press, 1990) p. 18.

[9] S. Shapiro and T. Kouri Kissel, Classical Logic, in *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta and U. Nodelman (Metaphysics Research Lab, Stanford University, 2024) Spring 2024 ed.

[10] G. Gentzen, Untersuchungen über das logische schließen. i, Mathematische Zeitschrift **39**, 176 (1935).

[11] F. J. Pelletier and A. Hazen, Natural Deduction Systems in Logic, in *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta and U. Nodelman (Metaphysics Research Lab, Stanford University, 2024) Spring 2024 ed.

[12] J. H. Sobel, On being and saying: Essays for richard cartwright (MIT Press, 1987) p. 241–261.

Appendix A: Utility Derivations

Lemma 1. *For any propositions, p and q,* $(p \to q) \to (\neg q \to \neg p)$

Proof.

$$\begin{prooftree}
\AxiomC{$[p \to q]^3$}
\AxiomC{$[p]^1$}
\RightLabel{$\to$ Elim}
\BinaryInfC{q}
\AxiomC{$[\neg q]^2$}
\RightLabel{$\neg$ Elim}
\BinaryInfC{$\bot$}
\LeftLabel{(1)}
\RightLabel{$\bot_c$}
\UnaryInfC{$\neg p$}
\LeftLabel{(2)}
\RightLabel{$\to$ Intro}
\UnaryInfC{$\neg q \to \neg p$}
\LeftLabel{(3)}
\RightLabel{$\to$ Intro}
\UnaryInfC{$(p \to q) \to (\neg q \to \neg p)$}
\RightLabel{$\forall$ Intro}
\UnaryInfC{$\forall p((p \to q) \to (\neg q \to \neg p))$}
\RightLabel{$\forall$ Intro}
\UnaryInfC{$\forall q \forall p((p \to q) \to (\neg q \to \neg p))$}
\end{prooftree}$$

□

Lemma 2. *For any model,* $M = \langle W, R, V \rangle$ *such that R is universal, for any proposition,q,* $\Diamond\Box q \to \Box q$

Proof. For any world, $\omega \in W$,

$$M, \omega \vDash \Diamond\Box q \iff \exists\omega' \in W \text{ s.t. } R\omega\omega',\ M, \omega' \vDash \Box q \tag{A1}$$

$$\iff \exists\omega' \in W\ \forall\omega'' \in W \text{ s.t. } R\omega\omega' \text{ and } R\omega'\omega'',\ M, \omega'' \vDash q. \tag{A2}$$

As R is universal, $\forall\omega, \omega', \omega'' \in W$, $R\omega\omega'$ and $R\omega'\omega''$ are true. Therefore:

$$M, \omega \vDash \Diamond\Box q \iff \forall\omega'' \in W,\ M, \omega'' \vDash q \iff \forall\omega'' \in W, \text{ s.t. } R\omega\omega'',\ M, \omega'' \vDash q \tag{A3}$$

$$\iff M, \omega \vDash \Box q. \tag{A4}$$

□