Basic Rules of Arithmetic

Julian J. Schloeder
Philosophy Department
University of Connecticut, Storrs

Forthcoming in *Australasian Journal of Philosophy*
(Preprint – Accepted Manuscript)

Abstract

Inferential expressivism makes a systematic distinction between inferences that are valid *qua* preserving commitment and inferences that are valid *qua* preserving evidence. I argue that the characteristic inferences licensed by the principle of comprehension, from x *is* P to x *is in the extension of* P and *vice versa*, fail to preserve evidence, but do preserve commitment. Taking this observation into account allows one to phrase inference rules for unrestricted comprehension without running into Russell's paradox. In the resulting logic, one can derive full second-order arithmetic. Thus, it is possible to derive classical arithmetic in a consistent logic with unrestricted comprehension.

Keywords: logicism, comprehension, Frege, Russell's paradox, inferential expressivism

1 Introduction

Naïve comprehension entails that for each formula there is its *extension*, containing all and only the things that satisfy the formula. There is obvious utility in being able to talk, *in toto*, about all the things that satisfy a formula, and in treating this totality as a singular object that can satisfy further formulas. It allows mathematicians to, for instance, speak of the totality of solutions to an equation and examine the properties of this totality. It allows the folk to speak of the properties of *collections*, as when one says that a museum's collection of artworks is remarkable.[1] Frege, albeit with hesitation, once laid down principles entailing comprehension to found arithmetic. Surely, comprehension ought not to be jettisoned lightly.

But there is nothing light about how naïve comprehension was jettisoned. It entails Russell's paradox. The standard diagnosis is that the self-referentiality allowed by naïve comprehension bears the blame, leading to the development of axiomatic set theory. Alternatively, one may revise the background logic to prevent (Weir, 1998; Field, 2008; Ripley, 2015) or tolerate the paradox (Restall, 1992; Weber, 2010). Neither approach gives much succor to Frege's (1884; 1893) foundational project. Frege's concerns put set-theoretic axioms equally on trial as they do number-theoretic axioms. And the known alternative background logics are too weak to derive classical arithmetic.[2] Neo-Fregeans, for their part, have given up on (unrestricted) comprehension (Hale and Wright, 2001).

[1]Plural logic is the attempt to regain these benefits without treating totalities as singular (Boolos, 1984; Oliver and Smiley, 2013). I set this to the side.

[2]Weber (2010), for instance, derives all *axioms* of arithmetic, but not all classical *theorems* of arithmetic.

My contribution here is twofold. First, a technical result: the (to my knowledge) first consistent calculus in which (i) all inferences between x *is* P and x *is in the extension of* P are valid, and (ii) all theorems of second-order Peano Arithmetic are provable. Second, a new diagnosis of Russell's paradox: the paradox is engendered by treating the principle of comprehension as preserving evidence whereas it only preserves commitment. The distinction between preservation of commitment and of evidence is an affordance of *inferential expressivism* (Incurvati and Schlöder, 2022, 2023b). The diagnosis motivates the calculus, but the calculus is interesting by itself and may also be motivated differently.

I begin with an exposition of inferential expressivism and the inferential expressivist diagnosis of the Liar paradox (Incurvati and Schlöder, 2023a). Using it as a blueprint, I present my diagnosis of Russell's paradox in Section 3 and go on to derive arithmetic in Section 4. I conclude in Section 5 by discussing where this leaves Frege's foundational project.

2 Inferential Expressivism

Traditional inferentialists claim that the meanings of (some) linguistic items are determined by which inferences a sentence containing the item can feature in (Gentzen, 1935; Prawitz, 1965; Dummett, 1991). For instance, inferentialists consider the standard natural deduction rules for conjunction as determining the meaning of conjunction.

Incurvati and Schlöder (2022), taking up observations going back to Frege (1879, 1919), Dummett (1973), and Rumfitt (2000), note that these rules are only valid if logic is *assertoric*. If logic is, for instance, *rejective*, they are invalid, as from rejecting *A and B* it does not follow that one rejects *A*. Thus, the received rules for conjunction are properly stated as follows.

- Asserting *A* and asserting *B* entails asserting *A and B*.
- Asserting *A and B* entails asserting *A* and asserting *B*.

But what is it to infer an assertion? Taking cues from expressivism, Incurvati and Schlöder claim that to assert a sentence is to express a belief and thereby to undertake a commitment to expressing this belief.[3] Such commitment to attitude expressions is what is preserved in inference. So, the conjunction rules state that someone who is committed to expressing belief towards *A* and towards *B* is also committed to expressing belief towards *A and B*, and *vice versa*. Commitments towards attitude expressions, in turn, are characterized as follows. Someone is committed to expressing an attitude if: when the issue is raised, they are conversationally obliged to express the attitude or retract an earlier commitment.[4]

Thereby, Incurvati and Schlöder (2022, 2023b) arrive at inferential expressivism, the view that meanings are determined by inferential relations between attitude expressions. The *conjunction rules* are then as follows, writing $+A$ for the expression of belief towards A.

$$(+\wedge\text{I.})\frac{+A \quad +B}{+A\wedge B} \qquad (+\wedge\text{E.}_1)\frac{+A\wedge B}{+A} \qquad (+\wedge\text{E.}_2)\frac{+A\wedge B}{+B}$$

It is straightforward to include attitudes other than belief. For instance, *bilateralists* claim that meanings are determined by conditions on both assertion and rejection (Smiley, 1996).

[3]This differs from Brandom's (1994) claim that asserting is undertaking commitments towards sentences. Here it is commitments towards attitude expressions towards sentences.

[4]It is *commitment* towards attitude expressions that is preserved rather than attitudes outright, as otherwise there is a clutter problem (Harman, 1986). It would be absurd to say that someone who expressed a belief is also expressing all beliefs that follow from it, as these may be arbitrarily many (Restall, 2005).

Inferential expressivists give the bilateral meaning of negation by the *negation rules*, writing $-A$ for the expression of disbelief towards A.

$$(+\neg\text{I.})\frac{-A}{+\neg A} \qquad (+\neg\text{E.})\frac{+\neg A}{-A} \qquad (-\neg\text{I.})\frac{+A}{-\neg A} \qquad (-\neg\text{E.})\frac{-\neg A}{+A}$$

That is, $(+\neg\text{I.})$ states that if one rejects A, one is committed to expressing belief towards the negation of A, so once this commitment is pointed out, one must actually express this belief or retract one's earlier rejection. The other rules are understood similarly. The bilateralist is not done, however. Expressions of belief and disbelief must *coordinate* as contraries, so bilateralists lay down the *Smileian reductio* rules.[5]

$$(\text{Rejection})\frac{+A \quad -A}{\perp} \qquad (\text{SR}_1)\frac{\begin{matrix}[+A]\\ \vdots \\ \perp\end{matrix}}{-A} \qquad (\text{SR}_2)\frac{\begin{matrix}[-A]\\ \vdots \\ \perp\end{matrix}}{+A}$$

If negation is defined by the negation rules and expressions of belief and disbelief are coordinated by Smileian *reductio*, then classical *reductio* and double negation elimination are valid in the logic under $+$ (Rumfitt, 2000).

Now, preserving commitments towards attitude expressions differs from the more familiar conception of inference as preserving *evidence*. Prawitz (2015) holds that "the aim of … inferences is to make assertions *justified*" (p. 71, his emphasis) and hence an "inference is … legitimate, if a subject who makes the inference and has evidence for its premisses thereby gets evidence for the conclusion" (p. 73). Similar claims have been made, notably, by Dummett (1978, 1991) and in recent discussions of epistemic modals (e.g. Schulz 2010) . Incurvati and Schlöder consider the evidence-preserving inferences to be a subset of the commitment-preserving ones. Following Prawitz, they claim that an inference preserves evidence if the inference can be used to justify expressing the attitude in its conclusion.[6]

However, Dickie (2010) argued that bilateralists cannot claim that inference preserves evidence, as rejections are evidentially unspecific. For instance, one may justify rejecting *Homer wrote the Iliad* with evidence for *Homer did not write the Iliad* but also with evidence for *Homer did not exist.* If the rejection is justified by evidence for *Homer did not exist*, Dickie points out, one cannot use this rejection to justify asserting *Homer did not write the Iliad.*[7] So the negation rules do not preserve evidence. Bilateralists might respond that the sign $-$ only denotes *strong rejections*, i.e. rejections justifiable by evidence that also justifies the corresponding negative assertion. This saves the negation rules, but renders Smileian *reductio* invalid, since inferences towards absurdity are also unspecific. Expressing belief towards *Homer wrote the Iliad* is absurd when one also expressed belief towards *Homer did not exist.* Smileian *reductio* hence

[5]The sign $\perp$ here stands for the act of announcing *Contradiction!* (Incurvati and Schlöder, 2023b, p230). This is a contentless speech act, to be distinguished from performing an assertion with contradictory content. One may assert a contradictory content without without realizing that this content is in fact contradictory. The function of announcing *Contradiction!* is to register an inconsistency; to be committed to $\perp$ is to be obliged to retract an earlier commitment. Also see Tennant 1999 for the related use of $\perp$ as a punctuation mark.

[6]Incurvati and Schlöder (2023a) stress that this is compatible with many conceptions of evidence, e.g. identifying evidence with rational credence (Schulz, 2010), or with knowledge (Williamson, 2000), but not with *any* conception. For instance, it rules out anti-luminosity (whereby the mere *existence* of an inference means one has evidence) as one cannot use evidence for justification if one does not know one has it.

[7]One might think that conceiving of negation as external negation or as metalinguistic negation helps, but neither ultimately meets Dickie's challenge (Schlöder, 2022).

entails that expressing belief towards *Homer did not exist* entails the rejection of *Homer wrote the Iliad*. If rejection is strong, this is mistaken.

Bilateralists can meet this challenge by becoming *multilateralists*. Reserving the sign $-$ for strong rejection (expressing disbelief), Incurvati and Schlöder (2022) add a third sign $\ominus$ for rejection *tout court*, expressing that one refrains from believing. Then the negation rules, phrased with the sign $-$ for strong rejection, remain valid as they preserve both commitment and evidence. And the Smileian *reductio* rules can be rephrased as follows to also preserve commitment and evidence.

$$(\text{Rejection})\,\frac{+A \qquad \ominus A}{\bot} \qquad (\text{SR}_1)\,\frac{\begin{matrix}[+A]\\ \vdots\\ \bot\end{matrix}}{\ominus A} \qquad (\text{SR}_2)\,\frac{\begin{matrix}[\ominus A]\\ \vdots\\ \bot\end{matrix}}{+A}$$

However, Smileian *reductio* was needed to ensure that $+$ and $-$ are contraries, which is not achieved by this version. Incurvati and Schlöder (2023a) suggest to recover valid versions of Smileian *reductio* for $+$ and $-$ by repeating the strategy used to recover valid versions of the negation rules. While the negation rules for rejections *tout court* are invalid, one can isolate a class of rejections, the strong ones, for which they are valid. Now, while the Smileian *reductio* rules for $+$ and $-$ are invalid for derivations of absurdity *tout court*, one can isolate a class of such derivations for which they are valid. This, Incurvati and Schlöder (2023a) argue, is just the class of evidence-preserving inferences, so they lay down the following *Smileian reductio** rules.

$$(\text{Rejection}^*)\,\frac{+A \qquad -A}{\bot} \qquad (\text{SR}_1^*)\,\frac{\begin{matrix}[+A]\\ \vdots\\ \bot\end{matrix}}{-A}\ \begin{matrix}\text{if the inference to } \bot\\ \text{preserves evidence}\end{matrix} \qquad (\text{SR}_2^*)\,\frac{\begin{matrix}[-A]\\ \vdots\\ \bot\end{matrix}}{+A}\ \begin{matrix}\text{if the inference to } \bot\\ \text{preserves evidence}\end{matrix}$$

Dickie's counterexamples are now ineffective. Inferring absurdity from *Homer did not exist* and *Homer wrote the Iliad* does not preserve evidence, since Homer existing is a precondition for there being evidence for *him* having written the Iliad. Thus if *Homer did not exist* is a premiss, *Homer wrote the Iliad* cannot occur in an evidence-preserving inference at all. It remains correct to infer absurdity from *Homer did not exist* and the assumption *Homer wrote the Iliad*. This inference preserves commitment and can properly feature in Smileian *reductio*. But it does not preserve evidence, so is excluded from Smileian *reductio** and one cannot infer the strong rejection of *Homer wrote the Iliad* on its basis.

Which inferences preserve evidence depends on the expressiveness of one's language. In the comparatively impoverished language of propositional logic, an inference can only fail to preserve evidence if it uses a premiss signed with $\ominus$. Thus, Incurvati and Schlöder (2023a) specify the following for propositional multilateral logic.

$$(\text{SR}_1^*)\,\frac{\begin{matrix}[+A]\\ \vdots\\ \bot\end{matrix}}{-A}\ \begin{matrix}\text{if no premisses signed with } \ominus\\ \text{were used to derive } \bot\end{matrix} \qquad (\text{SR}_2^*)\,\frac{\begin{matrix}[-A]\\ \vdots\\ \bot\end{matrix}}{+A}\ \begin{matrix}\text{if no premisses signed with } \ominus\\ \text{were used to derive } \bot\end{matrix}$$

The result is *basic multilateral logic* (BML). The *sentences* of BML are formed in the usual way from a countable set of propositional atoms, conjunction, and negation. A *formula* of BML is

$\perp$ or a sentence of BML prefixed with one of $+$, $-$ and $\ominus$. The calculus of BML is the natural deduction calculus over the negation rules (for $-$), the conjunction rules, Smileian *reductio* (for $+/\ominus$) and Smileian *reductio** (for $+/-$).

Incurvati and Schlöder (2023a) then lay down the following *truth rules* as determining the meaning of the truth predicate T, capturing its use as a device for expressing endorsement (Horwich, 1998) and opposition (Scharp, 2013).

$$(+\text{T-IN})\frac{+A}{+T\ulcorner A\urcorner}\qquad(+\text{T-OUT})\frac{+T\ulcorner A\urcorner}{+A}\qquad(-\text{T-IN})\frac{-A}{-T\ulcorner A\urcorner}\qquad(-\text{T-OUT})\frac{-T\ulcorner A\urcorner}{-A}$$

If one adds these rules (adjusting the language accordingly) to BML, paradox immediately follows. Following the usual presentation of the Liar paradox, consider a sentence L such that $\vdash +L \leftrightarrow \neg T\ulcorner L\urcorner$ and demonstrate that $+L \vdash +\neg L$ and $+\neg L \vdash +L$. If either of these inferences preserves evidence, one can apply Smileian *reductio** to derive the premiss of the other, and thus derive a contradiction.

But if either inference preserves evidence and it is possible at all to have evidence for L or $\neg L$, then the same evidence can justify both L and $\neg L$. One may say that it is fine for the same evidence to justify both a sentence and its negation, which is a version of the paraconsistent diagnosis (Priest, 1979). Or that it is not possible to have evidence for L or $\neg L$ at all, which is a version of the paracomplete diagnosis (Field, 2008). Incurvati and Schlöder (2023a) suggest a third option: that the truth rules do not preserve evidence, but are nonetheless valid *qua* preserving commitment. They contend that their diagnosis is on a par with the paracomplete or paraconsistent ones, as each is equally motivated by the Liar.

Incurvati and Schlöder (2023b, ch. 7) adduce another argument to support their diagnosis.[8] It begins with a tale told by Shapiro (2003). He imagines a disciple attending to a guru while a logician observes. The faithful disciple endorses everything asserted by the guru. The guru and the logician speak both the languages of arithmetic and set theory, but the disciple only speaks the language of arithmetic and is unable to understand the language of set theory. The guru already uttered some standard bridging principles between arithmetic and set theory and, thus, the disciple endorsed them. Now, the guru asserts a set-theoretic sentence A with no translation in the language of arithmetic. Despite not understanding A, the faithful disciple asserts $\ulcorner A\urcorner$ *is true*. The logician informs the disciple that $\ulcorner A\urcorner$ entails an arithmetical sentence $\ulcorner B\urcorner$, given the bridge principles. The disciple, taking this as competent testimony, expresses belief towards B.

Thus, one can intelligibly predicate truth of sentences one does not, and perhaps cannot, understand, mean something by it (i.e. express something one understands) and be tied to the consequences of this predication. Mundane examples of this abound, as when one says *Everything Aristotle said is true* without understanding Ancient Greek. The truth rules must hence apply to all sentences whatsoever. But then they cannot preserve evidence. A speaker who does not understand a sentence A may nevertheless understand $\ulcorner A\urcorner$ *is true* and can be justified in asserting $\ulcorner A\urcorner$ *is true*. But they cannot be justified in asserting A, since they do not understand A and hence cannot understand what would constitute evidence for A. It follows that the inference from $\ulcorner A\urcorner$ *is true* to A fails to preserve evidence. Since a single counterexample suffices to show that an inference rule does not preserve some property, the truth rules do not preserve evidence. However, one is bound by the consequences of one's truth predications, as highlighted by the role of the logician in Shapiro's tale. Thus, the truth rules preserve commitment.

[8]Incurvati and Schlöder (2023a) give yet another argument, but it is not relevant for what is to follow.

The argument rests on Incurvati and Schlöder's inferentialism: that to understand a sentence is to grasp its inferential role and that evidence is obtained by inference. Without grasping the inferential role of a sentence, one cannot apprehend inferences that conclude the sentence. Thus, one cannot have evidence for sentences one does not understand. These inferentialist claims are technical and might not entirely match folk uses of the term *evidence* (e.g. some might say that the guru's assertion is evidence for A regardless of whether A is understood).

Incurvati and Schlöder (2023a) conclude that the truth rules must be excluded from Smileian *reductio**. They demonstrate that, when this is done, the truth rules are consistent with BML.

3 Comprehension

Incurvati and Schlöder's diagnosis of the Liar can serve as a blueprint for diagnosing Russell's paradox. To begin, the following *conversion rules* define the meaning of λ-terms in a second-order logic.

$$(+\lambda\text{I.})\frac{+A[x/v]}{+[\lambda v.A]x} \qquad (+\lambda\text{E.})\frac{+[\lambda v.A]x}{+A[x/v]} \qquad (-\lambda\text{I.})\frac{-A[x/v]}{-[\lambda v.A]x} \qquad (-\lambda\text{E.})\frac{-[\lambda v.A]x}{-A[x/v]}$$

Comprehension can be expressed as the following inferences: for all objects x and concepts P, x is a member of its extension ϵP if and only if x is P. This corresponds to the following *extension rules*.[9]

$$(+\epsilon\text{I.})\frac{+Px}{+x \in \epsilon P} \qquad (+\epsilon\text{E.})\frac{+x \in \epsilon P}{+Px} \qquad (-\epsilon\text{I.})\frac{-Px}{-x \in \epsilon P} \qquad (-\epsilon\text{E.})\frac{-x \in \epsilon P}{-Px}$$

Extending BML with the conversion rules and the extension rules, without further changes, leads to Russell's paradox. Let R abbreviate $[\lambda v.\neg v \in v]$. Then $+\epsilon R \in \epsilon R \vdash +\neg\epsilon R \in \epsilon R$ and $+\neg\epsilon R \in \epsilon R \vdash +\epsilon R \in \epsilon R$, as witnessed by the following derivations.

$$+\epsilon R \in \epsilon R \vdash^{(+\epsilon\text{E.})} +[\lambda v.\neg v \in v](\epsilon R) \vdash^{(+\lambda\text{E.})} +\neg\epsilon R \in \epsilon R$$

$$+\neg\epsilon R \in \epsilon R \vdash^{(+\lambda\text{I.})} +[\lambda v.\neg v \in v](\epsilon R) \vdash^{(+\epsilon\text{I.})} +\epsilon R \in \epsilon R$$

If either of these inferences preserves evidence, Smileian *reductio** entails the premiss of the other, which then entails a contradiction. And if, additionally, it is at all possible to have evidence for either premiss, then the same evidence can justify both a sentence and its negation. This suggests the usual diagnoses. Paracompletists might claim that it is not possible to have evidence for $\epsilon R \in \epsilon R$ or $\neg\epsilon R \in \epsilon R$, whereas paraconsistentists might find it tolerable to have the same evidence supporting both sentences. My diagnosis is that the extension rules fail to preserve evidence. This is on a par with the paracomplete and paraconsistent diagnoses, but Shapiro's tale again yields an independent argument.

Consider a guru who makes pronouncements like x *is* P and a disciple who responds with x *is in* ϵP. The disciple can do this without understanding P. Perhaps the guru believes that $2^\omega = \aleph_2$ and asserts that 0 and 1 are the indices of $\aleph$-numbers below the continuum. The

[9] Technically, as relations also have extensions ('courses-of-values'), x abbreviates a sequence of n terms, n being the arity of P.

disciple cannot understand the alephs, but understands 0 and 1 and assents to the claim that they are in the extension of whatever property the guru talked about. This commits him to whatever arithmetical facts are entailed by $2^{\omega} = \aleph_2$. Thus, one can intelligibly utter sentences that contain, in transparent contexts, references to the extension of a concept that one does not understand, mean something by it (i.e. express something that one understands) and be tied to what follows from such utterances.

Again, mundane examples abound. Few know that *Platonic solid* means *convex, regular polyhedron in 3D space*, but some may nonetheless know that the cube is in the extension of *Platonic solid*. Someone in this epistemic situation can assert that the cube is in the extension of *Platonic solid*, mean something by it, and be thereby committed to the consequences of this claim. But, not understanding the concept, they do not grasp what would be evidence for *The cube is a Platonic solid* – they do not grasp that to justify this claim they must justify *The cube is regular and convex*. However, if they assert *The cube is in the extension of Platonic solid*, they are committed to *The cube is convex*, since they must assent to this (or retract) when it is pointed out to them that this is a consequence of their assertion.

These observations support an argument for the extension rules failing to preserve evidence. To stress, it relies on Incurvati and Schlöder's inferentialism: that to understand a sentence is to grasp its inferential role and that one cannot have evidence without grasping inferential role. As said, these claims might deviate from folk use of *evidence*. Moreover, the difference between extension-talk and concept-talk is often obscured. Someone who does not understand the concept *Platonic solid*, but believes that the cube is in its extension, might well assert *The cube is a Platonic solid*.

This is an imprecision of natural language. Perhaps all that someone knows about Platonic solids is that the list *tetrahedron, cube, octahedron, dodecahedron, icosahedron* exhausts them, so they use *Platonic solid* as a shorthand for this list. But there is an important difference between using a predicate as a shorthand for an extension (when one's understanding of *Platonic solid* is a list of five polyhedra) and as expressing an intension (when one's understanding of *Platonic solid* is *convex, regular polyhedron in 3D space*), despite either use giving rise to utterances like *The cube is a Platonic solid*. The distinction matters since there is a difference in inferential role. The claim *The cube is a Platonic solid*, understood intensionally, can be used in an inference to justify *The cube is convex*. But this is not so when it is understood extensionally, as then it is a claim about the cube being one of five polyhedra, from which nothing about convexity follows.[10]

Generally, having evidence for an object falling under a concept allows one to obtain, by inference, evidence for the object having the properties that make up the concept. Thus, if moving from extension to intension preserved evidence, then having evidence for an object belonging to a list (the extensional use) would allow one to obtain, by inference, evidence for it falling under a concept (the intensional use) and, thereby, obtain evidence for it having the properties that make up the concept. But this is not the case. Evidence for an object belonging to a list does not allow one to justify it having such properties, if one does not understand the

[10]This echoes Frege's (1892b) discussion of sense and reference. My distinction is between using *Platonic solid* with its sense *regular, convex polyhedron* and using it to directly name this sense's reference. Frege rejects direct naming, so might say that the latter use is also expressing a sense, namely *one of: tetrahedron, cube, octahedron, dodecahedron, icosahedron*. The argument proceeds the same either way.

Even in technical contexts, such imprecisions occur, e.g. when two mathematicians use *ordinal* to express different but equivalent definitions. Strictly speaking, moving from a claim about ordinals in one sense to a claim about ordinals in another sense is not evidence-preserving. This is ignored in practice, as the material equivalence of the senses is an understood theorem. Moving from one sense to the other then goes by *modus ponens* and the theorem, which does preserve evidence.

concept that gives rise to the list. As observed above, it is possible to have evidence for an object belonging to a list (e.g. from testimony) without understanding the concept that gives rise to the list. Nonetheless, such inferential moves preserve commitment, as someone who makes the extensional claim can be held to the consequences of the intensional claim, once pointed out.

In the formal language one makes the distinction sharp, reserving predication for the intensional use and membership/extension talk for the extensional use. But then the extension rules fail to preserve evidence. So, when extending BML with the conversion rules and the extension rules, we must exclude the latter from Smileian *reductio**.

$$\text{(SR}_1^*)\dfrac{\begin{array}{c}[+A]\\ \vdots\\ \bot\end{array}}{-A}\ \begin{array}{l}\text{if the inference to } \bot \text{ uses no premisses}\\ \text{signed with } \ominus \text{ and no extension rules}\end{array}$$

$$\text{(SR}_2^*)\dfrac{\begin{array}{c}[-A]\\ \vdots\\ \bot\end{array}}{+A}\ \begin{array}{l}\text{if the inference to } \bot \text{ uses no premisses}\\ \text{signed with } \ominus \text{ and no extension rules}\end{array}$$

Russell's paradox is then treated as follows. The derivations of $\bot$ from $+\epsilon R \in \epsilon R$ and from $+\neg\epsilon R \in \epsilon R$ are valid, but not evidence-preserving. Thus they cannot occur under Smileian *reductio**. But by Smileian *reductio* it follows that $\vdash \ominus\epsilon R \in \epsilon R$ and $\vdash \ominus\neg\epsilon R \in \epsilon R$. The claims that the extension of R is a member of itself and that it is not a member of itself are both to be rejected. Likewise for the claims that $R(\epsilon R)$ and $\neg R(\epsilon R)$.

However, there is a principled reason to reclaim some inferences as evidence-preserving. Consider two concepts P and Q and the complex λ-expression $[\lambda v.Pv \wedge Qv]$. To show that $x \in \epsilon[\lambda v.Pv \wedge Qv]$ entails $x \in \epsilon P$ requires the extension rules and hence does not preserve evidence. However, all that is required for evidence being preserved here is that conjunction is understood, not that P or Q is understood. Shapiro's guru might assert something about the conjunction of two set-theoretic concepts and, despite not understanding these concepts, the disciple can understand that the guru has asserted something *about a conjunction* and hence apprehend that if something is a member of the extension of the conjunctive concept, it is in the extension of either conjunct. So, in general, relations between concepts that can be grasped without understanding the concepts themselves, translate to the same relations between their extensions. Formally, this is expressed in the following meta-rule, where $\vdash^*$ denotes evidence-preserving inference.[11]

$$\text{(Absoluteness)}\dfrac{+P_1x_1, ..., +P_nx_n \vdash +Qy}{+x_1 \in \epsilon P_1, ..., +x_n \in \epsilon P_n \vdash +y \in \epsilon Q}$$

So if there is a derivation $+P_1x_1, ..., +P_nx_n \vdash^* +Qy$ (with no further premisses), there is a derivation $+x_1 \in \epsilon P_1, ..., +x_n \in \epsilon P_n \vdash^* +y \in \epsilon Q$. This is compatible with the guru/disciple argument because the argument showed that failures of evidence-preservation occur when one moves from extensional claims to intensional claims. But Absoluteness only adds inferences between extensional claims. Likewise, Absoluteness is compatible with the diagnosis of Russell's paradox, as the paradox essentially involves moves between extension and intension.

[11] Incurvati and Schlöder (2023a) discuss a similar rule for the truth predicate, but appear to consider it to be of mostly technical interest.

Now, the calculus BML$^\epsilon$ is defined as follows. Its language contains countably many first-order variables $x_0, x_1, \ldots$ and second-order variables $P_0, P_1, \ldots$ (with arities n_i), identity $=$, conjunction $\wedge$, negation $\neg$, first- and second-order universal quantifier $\forall$, 2-ary relation $\in$, term-forming operator ϵ, and concept-forming operator λ. Then define by simultaneous recursion:[12]

- A *term* is a first-order variable or ϵP where P is a predicate.
- A *predicate* of arity n is a second-order variable of arity n or $[\lambda x_1 \ldots x_n.A]$ where A is a sentence and the x_i are distinct first-order variables.
- An *atom* is one of $t_1 \ldots t_n \in t$, $t_1 = t_2$, or $Pt_1 \ldots t_n$ where t and the t_i are terms and P is an n-ary predicate.
- *Sentences* are defined in the usual way from atoms and $\neg$, $\wedge$ and $\forall$.

Then, the following are the obvious *quantifier rules*, where t in ($\forall$E.) ranges over arbitrary terms, and Q in ($\forall^2$E.) ranges over arbitrary predicates.

$$(+\forall\text{I.})\,\frac{+A[y/x]}{+\forall x(A)}\ \text{if } y \text{ does not occur free in premisses or undischarged assumptions} \qquad (+\forall\text{E.})\,\frac{+\forall x(A)}{+A[y/x]}$$

$$(+\forall^2\text{I.})\,\frac{+A[Q/P]}{+\forall P(A)}\ \text{if } Q \text{ does not occur free in premisses or undischarged assumptions} \qquad (+\forall^2\text{E.})\,\frac{+\forall P(A)}{+A[Q/P]}$$

The *identity rules* are as follows (cf. Read, 2004; Schlöder, 2023).

$$(+=\text{I.}^P)\,\frac{\begin{matrix}[+Px]\\ \vdots \\ +Py\end{matrix}}{+x=y}\ \text{where } P \text{ does not occur in premisses or assumptions} \qquad (+=\text{E.}^P)\,\frac{+Px \quad +x=y}{+Py}$$

$$(+=\text{I.}^{\epsilon P})\,\frac{\begin{matrix}[+x\in\epsilon P]\\ \vdots \\ +y\in\epsilon P\end{matrix}}{+x=y}\ \text{where } P \text{ does not occur in premisses or assumptions} \qquad (+=\text{E.}^{\epsilon P})\,\frac{+x\in\epsilon P \quad +x=y}{+y\in\epsilon P}$$

That is, two objects are identical if and only if they fall under all the same concepts or are in all the same extensions.

Let the calculus of BML$^\epsilon$ be the calculus of BML (with Smileian *reductio** amended to exclude the extension rules), extended with Absoluteness, conversion rules, extension rules, quantifier rules, and identity rules. Hereon, $\vdash$ is the calculus of BML$^\epsilon$. In the Appendix, I state a model theory for which it is sound.

Going forward, it will be useful to let $A \to B$ abbreviate $\neg(A \wedge \neg B)$, $A \vee B$ abbreviate $\neg(\neg A \wedge \neg B)$ and $\exists x(A)$ abbreviate $\neg\forall x(\neg A)$. For the conditional, one can derive *modus ponens* and, importantly, a *restricted version* of Conditional Proof (Incurvati and Schlöder, 2022).

[12] Someone worried about circularity or impredicativity here may find a stratified definition of such a language in Field et al. 2017, p454.

$$(+\to\text{I.})\dfrac{\begin{matrix}[+A]\\ \vdots \\ +B\end{matrix}}{+x=y}\ \text{if the inference to } \bot \text{ uses no premisses signed with } \ominus \text{ and no extension rules} \qquad (+\to\text{E.})\dfrac{+A\to B \quad +A}{+B}$$

4 Arithmetic

Zermelo (1908) defined the numbers by iteratively forming singletons, i.e. 0 is the empty set, 1 is $\{0\}$, 2 is $\{\{0\}\}$, *etc.* This requires very little set theory and is hence the most straightforward approach here.[13] So let S^x abbreviate $[\lambda v.v = x]$, and define $sx = \epsilon S^x$ and $0 = \epsilon[\lambda v.v \neq v]$.

A concept P is *inductive* if 0 falls under P and falling-under-P is closed under successor. Let Ω be the concept *falls under all inductive concepts.*

$$\Omega = [\lambda v.\forall P((P0 \wedge \forall y(Py \to Psy)) \to Pv)]$$

Trivially, Ω is inductive and $\Omega 0$, $\Omega s0$, *etc.* By the extension rules, $\omega = \epsilon\Omega$ exists and contains 0, $s0$, *etc.*[14] Define the concept *number* to be Ω and turn to the laws of arithmetic.

Theorem 4.1 (Induction)**.** $\vdash +\forall P((P0 \wedge \forall y(Py \to Psy)) \to \forall x(\Omega x \to Px))$

Proof. The derivation below shows t$+P0 \wedge \forall y(Py \to Psy), +\Omega x \vdash +Px$.

$$\dfrac{\dfrac{\dfrac{+\Omega x}{+\forall P((P0 \wedge \forall y(Py \to Psy)) \to Px)}}{+P0 \wedge \forall y(Py \to Psy)) \to Px}\,(+\forall^2\text{E.}) \qquad +P0 \wedge \forall y(Py \to Psy)}{+Px}\,(+\to\text{E.})$$

Everything in this derivation preserves evidence, so the theorem follows by $(+\to$I.) and $(+\forall$I.).

Next, derive the minimal structure of the number series: 0 is not a successor, successors are unique, and each number is either 0 or the successor of a number.

Theorem 4.2 (Robinson Axioms)**.** *The following are theorems of* BML^ϵ.

R1 $\forall x(\Omega x \to sx \neq 0)$.

R2 $\forall x\forall y((\Omega x \wedge \Omega y) \to (sx = sy \leftrightarrow x = y))$.

R3 $\forall x(\Omega x \to (x = 0 \vee \exists y(\Omega(y) \wedge x = sy)))$.

Proof of R3. Let $P = [\lambda v.v = 0 \vee \exists y(\Omega y \wedge v = sy)]$. By Induction, it suffices to show that $+P0$ and $+\forall n(Pn \to Psn)$. The base case, $+P0$, is trivial. So let n be arbitrary and suppose for the induction hypothesis that $+Pn$, i.e. $+n = 0 \vee \exists y(\Omega y \wedge n = sy)$. It is to show that $+sn = 0 \vee \exists y(\Omega y \wedge sn = sy)$.

Because Ω is inductive, $+\Omega 0$ and if $+\Omega y$, then $+\Omega sy$. So the induction hypothesis entails $+\Omega n$. Then, trivially, $+\Omega n \wedge sn = sn$. By existential generalization, $+\exists y(\Omega y \wedge sn = sy)$. This concludes the induction.

The proofs of R1 and R2 require more finesse. Begin with two lemmas.

[13] Von Neumann ordinals are superior to Zermelo's in that they extend to the transfinite and permit an easy definition of cardinality. But Zermelo ordinals suffice for arithmetic and are easier to handle.

[14] This allows one to identify the numbers, meeting Frege's requirement from the Julius Caesar Problem. To know whether Caesar is a number, one only needs to check whether Caesar falls under all inductive concepts.

Lemma 4.3. $\vdash +x \in sx$

Proof. Trivially, $(+=\mathrm{I.}^P)$ entails that $\vdash +x = x$ for any term x. So, by $(+\lambda\mathrm{I.})$, $\vdash +[\lambda v.v = x]x$ and hence, by $(+\epsilon\mathrm{I.})$, $\vdash +x \in sx$.

Lemma 4.4. $\vdash +\neg x \in 0$

Proof. As before, $\vdash +x = x$, so by $(-\neg\mathrm{I.})$, $\vdash -\neg x = x$, hence by $(-\lambda\mathrm{I.})$, $\vdash -[\lambda v.\neg v = v]x$. Thus, by $(-\epsilon\mathrm{I.})$, $\vdash -x \in \epsilon[\lambda v.\neg v = v]$, which by definition is just $\vdash -x \in 0$. Hence $+\neg x \in 0$ by $(+\neg\mathrm{I.})$.[15]

Proof of R1 Consider the following *pseudo*-derivation (superscripted 0 denotes an empty discharge).

$$
\begin{prooftree}
\AxiomC{Lm 4.3}
\UnaryInfC{$+x \in sx$}
\AxiomC{$[+sx = 0]^1$}
\RightLabel{$(+=\mathrm{E.}^{\epsilon P})$}
\BinaryInfC{$+x \in 0$}
\AxiomC{Lm 4.4}
\UnaryInfC{$+\neg x \in 0$}
\RightLabel{$(+\neg\mathrm{E.})$}
\UnaryInfC{$-x \in 0$}
\RightLabel{(Rejection*)}
\BinaryInfC{$\bot$}
\RightLabel{$(\mathrm{SR}_1^*)^1$}
\UnaryInfC{$-sx = 0$}
\RightLabel{$(+\neg\mathrm{I.})$}
\UnaryInfC{$+sx \neq 0$}
\RightLabel{$(+\to\mathrm{I.})^0$}
\UnaryInfC{$+\Omega x \to sx \neq 0$}
\RightLabel{$(+\forall\mathrm{I.})$}
\UnaryInfC{$+\forall x(\Omega x \to sx \neq 0)$}
\end{prooftree}
$$

Note that it is *not* a derivation because (SR_1^*) is applied to a subderivation that contains the lemmas, whose derivations do not preserve evidence. But this issue can be avoided by rewriting the pseudo-derivation to the following one.

$$
\begin{prooftree}
\AxiomC{$[+x \in sx]^4$}
\AxiomC{$[+sx = 0]^1$}
\RightLabel{$(+=\mathrm{E.}^{\epsilon P})$}
\BinaryInfC{$+x \in 0$}
\AxiomC{$[+\neg x \in 0]^2$}
\RightLabel{$(+\neg\mathrm{E.})$}
\UnaryInfC{$-x \in 0$}
\BinaryInfC{$\bot$}
\RightLabel{$(\mathrm{SR}_1^*)^1$}
\UnaryInfC{$-sx = 0$}
\RightLabel{$(+\neg\mathrm{I.})$}
\UnaryInfC{$+sx \neq 0$}
\RightLabel{$(+\to\mathrm{I.})^0$}
\UnaryInfC{$+\Omega x \to sx \neq 0$}
\RightLabel{$(+\forall\mathrm{I.})$}
\UnaryInfC{$+\forall x(\Omega x \to sx \neq 0)$}
\AxiomC{$[\ominus\forall x(\Omega x \to sx \neq 0)]^3$}
\BinaryInfC{$\bot$}
\RightLabel{$(\mathrm{SR}_1)^2$}
\UnaryInfC{$\ominus\neg x \in 0$}
\AxiomC{Lm 4.4}
\UnaryInfC{$+\neg x \in 0$}
\BinaryInfC{$\bot$}
\RightLabel{$(\mathrm{SR}_1)^4$}
\UnaryInfC{$\ominus x \in sx$}
\AxiomC{Lm 4.3}
\UnaryInfC{$+x \in sx$}
\BinaryInfC{$\bot$}
\RightLabel{$(\mathrm{SR}_1)^3$}
\UnaryInfC{$+\forall x(\Omega x \to sx \neq 0)$}
\end{prooftree}
$$

This *is* a derivation, since the lemmas do not occur under Smileian *reductio**.

The method used in this proof can be generalized to the *Theorem Theorem*, a meta-logical result about how theorems of BML^ϵ can be used in restricted proof contexts.[16]

Theorem 4.5 (Theorem Theorem). *If $\vdash +A$ and D is a pseudo-derivation that would be a derivation save that the proof of $+A$ occurs in a proof context restricted to evidence-preserving inference, then there is a derivation D' with the same premisses and conclusion as D.*

Proof. Suppose we have such a pseudo-derivation D of $\Gamma \vdash +C$. Schematically:

[15] This proof highlights the need for the extension rules for $-$. They are how one shows that something is *not* in an extension.

[16] Incurvati and Schlöder (2023a) discuss a special case of this result.

$$\dfrac{\overline{+A} \qquad \Gamma}{\begin{array}{c} D \\ +C \end{array}}$$

Such pseudo-derivations can be rewritten to proper derivations as follows.

$$\dfrac{\dfrac{\dfrac{\dfrac{[+A]^1 \qquad \Gamma}{\begin{array}{c} D \\ +C \end{array}} \qquad [\ominus C]^2}{\bot}}{\ominus A}\,(\mathrm{SR}_1)^1 \qquad \overline{+A}}{\dfrac{\bot}{+C}\,(\mathrm{SR}_2)^2}$$

The Theorem Theorem states that if $+A$ as a theorem, it does not matter whether its proof uses rules that do not preserve evidence. Loosely said, being committed to $+A$ as a matter of logic (i.e. from no assumptions) is evidence for $+A$.

Proof of R2. For any P, the following derivation witnesses $+S^y x, +Px \vdash^* +Py$. This is because $+S^y x$ is $+[\lambda v.v = y]x$ which by ($+\lambda$E.) entails $+x = y$. By ($+ =$E.^{P}), this and $+Px$ entails $+Py$.

By Absoluteness, $+x \in sy, +x \in \epsilon P \vdash^* +y \in \epsilon P$. Thus, by ($+ =$I.$^{\epsilon P}$), $+x \in sy, \vdash +x = y$. By Lemma 4.3, $\vdash +x \in sx$, so by ($+ =$E.$^{\epsilon P}$), $+sx = sy \vdash +x \in sy$. Hence, $+sx = sy \vdash +x = y$. By Theorem 4.5, we may apply ($+ \to$I.), so $\vdash +(sx = sy \to x = y)$. The converse follows immediately from the identity rules.

This concludes the proof of Theorem 4.2. As is known, R1–R3 plus Induction entail the Dedekind–Peano Axioms in second-order logic.[17] Since BML^ϵ extends second-order logic and by the Theorem Theorem one may treat R1–R3 as if they were axioms, it follows that BML^ϵ includes second-order arithmetic. One can appeal to R1–R3 and Induction while remaining in the evidence-preserving fragment of BML^ϵ, so full Conditional Proof and classical *reductio* are available. Thus, every proof in second-order arithmetic can be expressed in BML^ϵ.

Theorem 4.6. BML^ϵ *includes second order Peano Arithmetic.*

One can now define $n < m$ iff $\exists k(\Omega k \wedge k \neq 0 \wedge n + k = m)$ and define *number-of* as $\#P = n$ iff $\Omega n \wedge P \approx [\lambda x.x < n]$, where $\approx$ is the usual second-order definition of equinumerosity. Note that this assigns a number only to concepts with finite extensions.[18]

5 Foundations

In the preface to the *Grundgesetze*, Frege remarked that logicians routinely rely on talk of extensions – later citing Leibniz and Boole as examples in §8 – and have merely failed to articulate the laws governing this talk. He hence considers such talk to be an affordance of logic itself and its laws to be laws of logic. Insofar as Frege had a definitive conception of extension, it was surely centered on the idea that extensions, unlike concepts, are objects, and

[17] This is because the following defines addition so that the usual recursive axioms follow by induction: $n+m = p$ iff $\forall F((F(0) = n \wedge \forall k(\Omega k \to (F(sk) = sF(k)))) \to F(m) = p)$. Multiplication is analogous.

[18] One can extend the definition by piecemeal, e.g. $\#P = \omega$ if P is equinumerous to Ω, and $\#P = 2^\omega$ if P is equinumerous to $[\lambda v.\forall x(x \in v \to \Omega x)]$, *etc*, but this yields no definition of transfinite number.

that besides the syntactic differences this entails (cf. Frege, 1892a), talk of a concept and talk of its extension are interchangeable.[19]

If we take my extension rules to represent the meaning-conferring inferential practices of extension talk, we may regard them as analytic, i.e. valid in virtue of what one means by such talk, and therefore, arguably, as candidates for logical laws. We may also regard them as explicating a Fregean conception of extension, as they allow unrestricted transitions between extension-talk and concept-talk. By taking a particular stance on the nature of meaning – the inferentialist one – one can recognize an important distinction in the sense *in which* the extension rules are valid in virtue of meaning. Namely, that they are valid in the sense of preserving commitment rather than in the sense of preserving evidence. The addition of Absoluteness does not complicate this picture, as it is motivated by recognizing the same distinction in the same practices.

This may appear to be a radical departure from Frege. He claimed that the logical laws are truths about real logical entities (such as extensions) and that analytic truths are simply those that can be derived from the logical laws and definitions. But the appearance may be misleading. Incurvati and Schlöder (2023b, ch. 7) argue that although their inferential expressivism does not require or entail objectively real entities, it also does not rule them out. In particular, they contend that one may claim that the meaning-conferring rules are "latching onto" real properties of real entities in the sense that they have become part of our languages in response to such properties and entities. So they might reconcile with Frege by simply *adding* a metaphysical claim to their semantic view; for whether Frege may reconcile with inferentialism, see Dummett 1973.[20]

One can, however, assess Frege's logicism while bracketing these matters. Logicism can be understood as both a metaphysical and an epistemological claim: that numbers are logical objects and that logic is how one learns truths about numbers. Put differently, logicism is the attempt to reduce questions about the metaphysics and epistemology of numbers to the metaphysics and epistemology of logic. Bracketing the questions about logic, we may ask whether inferential expressivism can make good on the reduction.

Frege began, in the *Grundlagen*, with a discussion of the *number-of* operation and its linguistic properties. He developed a formal definition of number on this basis, demonstrated the existence of a series of such numbers, derived the laws of arithmetic, and showed how to understand number-of talk in terms of this series. Thus, if Frege's laws are laws of logic, then logic entails that the number series exists and logical proof is how one learns truths about numbers. I defined numbers differently, but immaterially so. Zermelo's definition is motivated by similar (if somewhat more abstract) observations about number talk being about a series with a particular structure. Like Frege, I proceed by demonstrating that a series of such numbers exists, deriving the laws of arithmetic, and showing how to define *number-of* in terms

[19]In a famously confounding footnote in the *Grundlagen* (p. 80), Frege claims that extensions need no explanation, but also suggests to replace talk of extensions with talk of concepts altogether; specifically, that instead of defining numbers as extensions, one may define them as the corresponding concepts. This is in sharp contrast to the postface of the *Grundgesetze*, where he laments that he sees no way to define numbers if not as extensions. Frege's changing understanding is visible in the intermediate works where he claims that extensions are objects (1891) and that concepts cannot be objects (1892a). By the time of the *Grundgesetze* he may conceive of extensions as how logicians approximate talk of concepts as objects and consider such approximations necessary for logic, whence the logical objects. Cf. Klement's (2012) stronger claim that Frege conceives of extensions as 'nothing but [a] concept itself considered as an object'.

[20]Another major divergence is that BML^{ϵ} allows non-identical extensions with the same members, which is ruled out by Basic Law V. Extensionality is not needed for deriving arithmetic and its inclusion would complicate the construction in the Appendix, so I omit it. But one could add extensionality and (likely) address the complications; see Field et al. 2017 for extensionality in a similar construction.

of this series. So BML^ϵ entails that the number series exists and BML^ϵ inference is how one learns truths about numbers.

Thus, if Frege is right that talk of totalities in singular terms is part of logic and I am right that BML^ϵ captures such talk, then BML^ϵ vindicates the reductive ambitions of logicism. These are big *if*'s, of course. The extension rules are not ontologically neutral, putting pressure on their supposed logicality and analyticity (cf. Boolos, 1997). Frege considered logic to come with logical objects and, supposedly, disagrees with any conception of logicality as ontologically neutral. Inferential expressivists, Incurvati and Schlöder (2023b) argue, can respond to ontological challenges with either deflationary moves or additional metaphysical claims about these rules latching onto a prior reality. This allows them to treat rules like the extension rules as analytic. But Incurvati and Schlöder (2023b) provide no criteria of logicality.

Thus, inferential expressivism vindicates the *technical* ambition of the *Grundgesetze*, namely the reduction of arithmetic to comprehension-talk. To vindicate the reduction to *logic*, inferential expressivists must take a further step. They may determine some parts of inferential practice to be distinctively *logical* inferential practice and take Frege's comments about the use of extensions to show that comprehension-talk belongs to the logical practice, hence to logic. Or they may take additional assumptions about the metaphysics of logic on board and determine that the comprehension rules have latched onto this metaphysics. Both options would vindicate Frege's reduction of the metaphysics and epistemology of numbers to the metaphysics and epistemology of logic. And if inferential expressivists ultimately conclude that there are no such principled distinctions within inferential practice, then the reduction of arithmetic to comprehension-talk by itself remains a substantive thesis about the metaphysics and epistemology of numbers.

Either way, the rules of BML^ϵ are *Grundregeln*, basic rules, of arithmetic.[21]

Appendix

An *abstract* is a term of the form $\epsilon[\lambda x_1 ... x_n.A]$ where A is a sentence. The arity of an abstract is the number of distinct variables bound by λ. Let D be the set of all abstracts. An *interpretation* I assigns to each first-order variable a member of D and to each second-order variable with arity n a subset of D^n (where D^1 is D). Let $\sigma : D \to \bigcup_{n>0} \mathcal{P}(D^n)$ assign to each abstract with arity n a subset of D^n. Intuitively, σ assigns to each abstract its extension.

The interpretation of terms is as follows. $I(t) = t$ if t is an abstract, $I(t) = I(x)$ if t is a first order variable x, and $I(t) = [\lambda v_1 ... v_n.Pv_1 ... v_n]$ if $t = \epsilon P$ for a second order variable P with arity n. Satisfaction is defined as follows.

- $I, \sigma \models t_1 ... t_n \in t$ iff n is the arity of $I(t)$ and $(I(t_1), ..., I(t_n)) \in \sigma(I(t))$.
- $I, \sigma \models Pt_1 ... t_n$ iff n is the arity of P and $(I(t_1), ..., I(t_n)) \in I(P)$
- $I, \sigma \models t_1 = t_2$ iff $I(t_1) = I(t_2)$.
- $I, \sigma \models A \wedge B$ iff $I, \sigma \models A$ and $I, \sigma \models B$.
- $I, \sigma \models \neg A$ iff not $I, \sigma \models A$.
- $I, \sigma \models [\lambda v_1 ... v_n.A]t_1 ... t_n$ iff $I, \sigma \models A[t_1/v_1, ..., t_n/v_n]$.

[21] I am grateful to Marcus Rossberg, Stewart Shapiro, and the audience of the EuPhilo 2nd Annual Conference, held at the University of Bonn in 2023, for comments on earlier versions of this work.

- $I, \sigma \models \forall x A$ iff $I, \sigma \models A[t/x]$ for all terms t.
- $I, \sigma \models \forall P A$ iff $I, \sigma \models A[[\lambda v_1 ... v_n.B]/P]$ for all sentences B, where n is the arity of P.

The task is to find the right σ, which will go by a supervaluation over functions τ.

Definition 1. $\tau : D \to \bigcup_{n>0} \mathcal{P}(D^n)$ is *coherent* if for all sentences A_i:

1. $\tau(\epsilon[\lambda v_1 ... v_n.\neg A_1]) = D^n \setminus \tau(\epsilon[\lambda v_1 ... v_n.A_1])$.
2. $\tau(\epsilon[\lambda v.A_1 \wedge A_2]) = \tau(\epsilon[\lambda v.A_1]) \cap \tau(\epsilon[\lambda v.A_2])$.
3. $\tau(\epsilon[\lambda v_1 ... v_n.A_1 \wedge ... \wedge A_n]) = \tau(\epsilon[\lambda v_1.A_1]) \times ... \times \tau(\epsilon[\lambda v_n.A_n])$ if v_i is not free in A_j whenever $i \neq j$.

Let I be any interpretation. For $\tau, \sigma : D \to \bigcup_{n>0} \mathcal{P}(D^n)$ define $\tau \sqsupseteq \sigma$ iff for all abstracts a, $\tau(a) \supseteq \sigma(a)$. Then define a sequence σ_α.

- Base: $\sigma_0(a) = \emptyset$ for all a.
- Successor: for each $\epsilon[\lambda v_1 ... v_n.A] \in D$, let $(a_1, ..., a_n) \in \sigma_{\alpha+1}(\epsilon[\lambda v_1 ... v_n.A])$ iff for all coherent $\tau \sqsupseteq \sigma_\alpha$ it is the case that $I, \tau \models [\lambda v_1 ... v_n.A]a_1 ... a_n$.
- Limit: if λ is a limit, $\sigma_\lambda(a) = \bigcup_{\alpha<\lambda} \sigma_\alpha(a)$ for all a.

The σ_α sequence is non-decreasing, so there is a fixed point σ^I. Then define:

- $I \models +A$ iff for all coherent $\tau \sqsupseteq \sigma^I$, we have $I, \tau \models A$,
- $I \models -A$ iff for all coherent $\tau \sqsupseteq \sigma^I$, we have $I, \tau \models \neg A$,
- $I \models \ominus A$ iff there is a coherent $\tau \sqsupseteq \sigma^I$ such that $I, \tau \models \neg A$.
- never $I \models \bot$

Write $\Gamma \models \varphi$ iff for all I, if $I \models \psi$ for all $\psi \in \Gamma$, then $I \models \varphi$.

Theorem 5.1 (Soundness). *If $\Gamma \vdash \varphi$, then $\Gamma \models \varphi$.*

Proof. By induction on the construction of proofs $\mathcal{D}$.

The rules of BML are as in Incurvati and Schlöder 2023a, the identity rules are trivial, quantifiers are standard, and the conversion rules are coded in the model theory. It suffices to check the extension rules and Absoluteness.

Assume without loss of generality that P is of the form $[\lambda v.A]$ and t is an abstract. The cases where P is instead a second-order variable only requires minor modifications and the cases where P has arity $n > 1$ are analogous. The cases where t is x or ϵQ (where x is a first-order variable and Q is a second-order variable) only require replacing t with $I(t)$.

- ($+\epsilon$I.). Suppose $\mathcal{D}$ ends with $+Pt \vdash +t \in \epsilon P$. The induction hypothesis is $I \models +[\lambda v.A]t$. By definition, for all coherent $\tau \sqsupseteq \sigma^I$, $I, \tau \models [\lambda v.A]t$. As σ^I is a fixed point, $t \in \sigma^I(\epsilon[\lambda v.A])$. So for every $\tau \sqsupseteq \sigma^I$, $t \in \sigma^I(\epsilon[\lambda v.A])$. So for each such τ, by definition, $I, \tau \models t \in \epsilon[\lambda v.A]$. By definition, $I \models +t \in \epsilon[\lambda v.A]$.

- ($+\epsilon$E.). Suppose $\mathcal{D}$ ends with $+t \in \epsilon P \vdash +Pt$. The induction hypothesis is $I \models +t \in \epsilon[\lambda v.A]$. By definition, for all coherent $\tau \sqsupseteq \sigma^I$, $I, \tau \models t \in \epsilon[\lambda v.A]$. Hence for all such τ, it is the case that $t \in \tau(\epsilon[\lambda v.A])$. If it were not the case that $t \in \sigma^I(\epsilon[\lambda v.A])$ there would be a τ where this is not so. So $t \in \sigma^I(\epsilon[\lambda v.A])$. Because σ^I is a fixed point, for all coherent τ, $I, \tau \models [\lambda v.A]t$. So, by definition, $I \models +[\lambda v.A]t$.

- ($-\epsilon$I.). Suppose $\mathcal{D}$ ends with $-Pt \vdash -t \in \epsilon P$. The induction hypothesis is $I \models -[\lambda v.A]t$. By definition, for all coherent $\tau \sqsupseteq \sigma^I$, $I, \tau \models \neg[\lambda v.A]t$. This is $I, \tau \models [\lambda v.\neg A]t$. As σ^I is a fixed point, $t \in \sigma^I(\epsilon[\lambda v.\neg A])$. So for every $\tau \sqsupseteq \sigma^I$, $t \in \tau(\epsilon[\lambda v.\neg A])$. Since the τ are coherent, $t \notin \tau(\epsilon[\lambda v.A])$ So by definition $I, \tau \models \neg t \in \epsilon[\lambda v.A]$ for all such τ. By definition, $I \models -t \in \epsilon[\lambda v.A]$.

- ($-\epsilon$E.). Suppose $\mathcal{D}$ ends with $-t \in \epsilon P \vdash -Pt$. The induction hypothesis is $I \models -t \in \epsilon[\lambda v.A]$. By definition, for all coherent $\tau \sqsupseteq \sigma^I$, $I, \tau \models \neg t \in \epsilon[\lambda v.A]$. So for all such τ, it is the case that $t \notin \tau(\epsilon[\lambda v.A])$. As the τ are coherent, $t \in \tau(\epsilon[\lambda v.\neg A])$. If it were the case that $t \notin \sigma^I(\epsilon[\lambda v.\neg A])$ there would be a τ where $t \notin \tau(\epsilon[\lambda v.\neg A])$. Since this is not so, $t \in \sigma^I(\epsilon[\lambda v.\neg A])$. Because σ^I is a fixed point, for all coherent τ, $I, \tau \models [\lambda v.\neg A]t$. This is $I, \tau \models \neg[\lambda v.A]t$. So, by definition, $I \models -[\lambda v.A]t$.

Turn to Absoluteness. Because the sequents in Absoluteness are evidence-preserving, we may write them as material conditionals. So assume:

$$\models +\neg([\lambda v_1.A_1]t_1 \wedge ... \wedge [\lambda v_n.A_n]t_n \wedge \neg[\lambda v.A]t)$$

And show that:

$$\models +\neg(t_1 \in \epsilon[\lambda v_1.A_1] \wedge ... \wedge t_n \in \epsilon[\lambda v_n.A_n] \wedge \neg t \in \epsilon[\lambda v.A])$$

The assumption means that for all I and all coherent $\tau \sqsupseteq \sigma^I$:

$$I, \tau \models \neg([\lambda v_1.A_1]t_1 \wedge ... \wedge [\lambda v_n.A_n]t_n \wedge \neg[\lambda v.A]t)$$

That is:

$$I, \tau \models [\lambda v_1...v_n v.\neg(A_1 \wedge ... \wedge A_n \wedge \neg A)]t_1...t_n t$$

Then, since σ^I is a fixed point:

$$(t_1, ..., t_n, t) \in \sigma^I(\epsilon[\lambda v_1...v_n v.\neg(A_1 \wedge ... \wedge A_n \wedge \neg A)])$$

So for all coherent $\tau \sqsupseteq \sigma^I$:

$$(t_1, ..., t_n, t) \notin \tau(\epsilon[\lambda v_1...v_n v.(A_1 \wedge ... \wedge A_n \wedge \neg A)]) \qquad (*)$$

Now, assume for *reductio* that there is a coherent τ such that :

$$I, \tau \not\models \neg(t_1 \in \epsilon[\lambda v_1.A_1] \wedge ... \wedge t_n \in \epsilon[\lambda v_n.A_n] \wedge \neg t \in \epsilon[\lambda v.A])$$

That is:

$$I, \tau \models t_1 \in \epsilon[\lambda v_1.A_1] \wedge ... \wedge t_n \in \epsilon[\lambda v_n.A_n] \wedge \neg t \in \epsilon[\lambda v.A]$$

So for all $i \leq n$, $t_i \in \tau(\epsilon[\lambda v_i.A_i])$ and $t \notin \tau(\epsilon[\lambda v.A])$. τ is coherent, so $t \in \tau(\epsilon[\lambda v.\neg A])$ and so $(t_1, ..., t_n, t) \in \tau(\epsilon[\lambda v_1...v_n.(A_1 \wedge ... \wedge A_n \wedge \neg A)])$. This contradicts (*).

References

Boolos, George (1984) 'To Be is To Be a Value of a Variable (or To Be Some Values of Some Variables)'. *Journal of Philosophy* **81**: 430–449.

Boolos, George (1997) 'Is Hume's Principle analytic?' In Richard Kimberly Heck, ed., *Language Thought and Logic: Essays in Honour of Michael Dummett*, 245–261. Oxford: Clarendon.

Brandom, Robert (1994) *Making it Explicit*. Cambridge, MA: Harvard University Press.

Dickie, Imogen (2010) 'Negation, Anti-realism, and the Denial Defence'. *Philosophical Studies* **150**: 161–85. doi:10.1007/s11098-009-9364-z.

Dummett, Michael (1973) *Frege: Philosophy of Language*. London: Duckworth.

Dummett, Michael (1978) *Truth and Other Enigmas*. Cambridge, MA: Harvard University Press.

Dummett, Michael (1991) *The Logical Basis of Metaphysics*. Cambridge, MA: Harvard University Press.

Field, Hartry (2008) *Saving Truth From Paradox*. New York: Oxford University Press.

Field, Hartry, Harvey Lederman, and Tore Fjetland Øgaard (2017) 'Prospects for a Naive Theory of Classes'. *Notre Dame Journal of Formal Logic* **58**: 461–506. doi:10.1215/00294527-2017-0010.

Frege, Gottlob (1879) *Begriffsschrift: Eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle: Louis Nebert.

Frege, Gottlob (1884) *Die Grundlagen der Arithmetik: Eine logisch-mathematische Untersuchung über den Begriff der Zahl*. Breslau: Wilhelm Koebner.

Frege, Gottlob (1891) 'Funktion und Begriff'. In G Patzig, ed., *Funktion, Begriff, Bedeutung (1962)*, 1–22. Göttingen: Vandenhoeck & Ruprecht.

Frege, Gottlob (1892a) 'Über Begriff und Gegenstand'. *Vierteljahrsschrift für wissenschaftliche Philosophie* **16**: 192–205.

Frege, Gottlob (1892b) 'Über Sinn und Bedeutung'. *Zeitschrift für Philosophie und Philosophische Kritik* **100**: 25–50.

Frege, Gottlob (1893) *Grundgesetze der Arithmetik*, volume 1. Jena: Hermann Pohle.

Frege, Gottlob (1919) 'Die Verneinung: Eine logische Untersuchung'. *Beiträge zur Philosophie des deutschen Idealismus* **1**: 143–157.

Gentzen, Gerhard (1935) 'Untersuchungen über das logische Schließen I'. *Mathematische Zeitschrift* **39**: 176–210.

Hale, Bob and Crispin Wright (2001) *The Reason's Proper Study: Essays towards a Neo-Fregean Philosophy of Mathematics*. Oxford: Clarendon Press.

Harman, Gilbert (1986) *Change in View*. Cambridge, MA: MIT Press.

Horwich, Paul (1998) *Truth*. Oxford: Clarendon Press.

Incurvati, Luca and Julian J Schlöder (2022) 'Epistemic Multilateral Logic'. *Review of Symbolic Logic* **15**: 505–36. doi:10.1017/s1755020320000313.

Incurvati, Luca and Julian J Schlöder (2023a) 'Inferential Deflationism'. *Philosophical Review* 529–578. doi:10.1215/00318108-10697531.

Incurvati, Luca and Julian J Schlöder (2023b) *Reasoning With Attitude*. Oxford University Press.

Klement, Kevin C (2012) 'Frege's Changing Conception of Number'. *Theoria* **78**: 146–167. doi:10.1111/j.1755-2567.2012.01129.x.

Oliver, Alex and Timothy Smiley (2013) *Plural Logic*. Oxford University Press.

Prawitz, Dag (1965) *Natural Deduction. A Proof-Theoretical Study*. Uppsala: Almqvist & Wicksell.

Prawitz, Dag (2015) 'Explaining Deductive Inference'. In Heinrich Wansing, ed., *Dag Prawitz on Proofs and Meaning*, 65–100. Dordrecht: Springer.

Priest, Graham (1979) 'The Logic of Paradox'. *Journal of Philosophical logic* **8**: 219–41.

Read, Stephen (2004) 'Identity and Harmony'. *Analysis* **64**: 113–119. doi:10.1093/analys/64.2.113.

Restall, Greg (1992) 'A Note on Naive Set Theory in LP'. *Notre Dame Journal of Formal Logic* **33**: 422–432. doi:10.1305/ndjfl/1093634406.

Restall, Greg (2005) 'Multiple Conclusions'. In Petr Hájek, Luis Valdés-Villanueva, and Dag Westerståhl, eds., *Logic, Methodology and Philosophy of Science: Proceedings of the Twelfth International Congress*, 189–205. London: King's College Publications.

Ripley, David (2015) 'Naive Set Theory and Nontransitive Logic'. *The Review of Symbolic Logic* **8**: 553–571. doi:10.1017/s1755020314000501.

Rumfitt, Ian (2000) "'Yes" and "No"'. *Mind* **109**: 781–823. doi:10.1093/mind/109.436.781.

Scharp, Kevin (2013) *Replacing Truth*. Oxford: Oxford University Press.

Schlöder, Julian J (2022) 'Assertion and Rejection'. In Daniel Altshuler, ed., *Linguistics Meets Philosophy*, 414–438. Cambridge University Press.

Schlöder, Julian J (2023) 'Identity and Harmony and Modality'. *Journal of Philosophical Logic* **52**: 1269–1294. doi:10.1007/s10992-023-09705-8.

Schulz, Moritz (2010) 'Epistemic Modals and Informational Consequence'. *Synthese* **174**: 385–95. doi:10.1007/s11229-009-9461-8.

Shapiro, Stewart (2003) 'The Guru, the Logician, and the Deflationist: Truth and Logical Consequence'. *Noûs* **37**: 113–132. doi:10.1111/1468-0068.00431.

Smiley, Timothy (1996) 'Rejection'. *Analysis* **56**: 1–9. doi:10.1111/j.0003-2638.1996.00001.x.

Tennant, Neil (1999) 'Negation, Absurdity and Contrariety'. In D Gabbay and H Wansing, eds., *What is Negation?*, 199–222. Dordrecht: Kluwer.

Weber, Zach (2010) 'Transfinite Numbers in Paraconsistent Set Theory'. *Review of Symbolic Logic* **3**: 71–92. doi:10.1017/s1755020309990281.

Weir, Alan (1998) 'Naïve Set Theory is Innocent!' *Mind* **107**: 763–798. doi: 10.1093/mind/107.428.763.

Williamson, Timothy (2000) *Knowledge and Its Limits.* Oxford: Oxford University Press.

Zermelo, Ernst (1908) 'Untersuchungen über die Grundlagen der Mengenlehre. I.' *Mathematische Annalen* **65**: 261–281.