Natural constructive proofs of A via A -> B, proof paradoxes, and impredicativity

Mark van Atten

Chapter 1
Natural constructive proofs of A via $A \to B$, proof paradoxes, and impredicativity

Mark van Atten, March 6, 2022

for Göran

Abstract Guided by a passage in Kreisel, this is a discussion of the relations between the phenomena in the title, with special attention to the method of analysis and synthesis in Greek geometry, fixed point theorems, and Kreisel's contact with Gödel.

Key words: analysis and synthesis, Brouwer, constructivity, informal proof, fixed points, formal proof, Gödel, Goodman, Greek geometry, Heyting, implication, impredicativity, intuitionistic logic, Kreisel, Lawvere's Fixed Point Theorem, natural deduction, paradoxes, proof, Proof Explanation, Troelstra

Mark van Atten
Archives Husserl, Pays Germaniques (CNRS / ENS), 45 rue d'Ulm, 75005 Paris, France, e-mail: mvanatten@ens.fr

1.1 Introduction: a passage in Kreisel

In his obituary for Brouwer in the *Memoirs of the Royal Society*, Kreisel made this parenthetical remark:

(1.1) (Incidentally, it is one of the peculiarities of constructive logic that, for some A, a *natural* formal proof of A goes *via* proofs of $A \rightarrow B$ and of $(A \rightarrow B) \rightarrow A$: such a proof of A actually contains a proof of $A \rightarrow B$.) (Kreisel & Newman, 1969, p. 58, original emphasis)

My aim is to analyse this remark and its role in the immediate context in which it was made. This introduction provides that context, presents the questions that will be treated, and comments on the structure of the paper.

Earlier on in the memoir, Kreisel explains the meaning of constructive implication thus:

(1.2) to prove $A \rightarrow B$, one needs two things: a mapping π from proofs to proofs, and a proof, say p_0, establishing that, if any p proves A then $\pi(p)$ proves B. (Kreisel & Newman, 1969, p. 57)

That is his construal of the corresponding clause in Heyting's Proof Explanation:

(1.3) The implication $p \rightarrow q$ can be asserted, if and only if we possess a construction r, which, joined to any construction proving p (supposing that the latter be effected), would automatically effect a construction proving q. (Heyting, 1956, p. 98)

Kreisel, keeping (1.2) in mind, a bit further on writes:

(1.4a) Perhaps because of all this experience or for intrinsic reasons, nobody seems ever to have been as much as tempted to put down false principles in elementary constructivity. In contrast, if one actually wants to formulate explicit properties of proofs, one has to keep one's wits about one to avoid errors which are, formally, similar to Russell's paradox in set theory.

(1.4b) This is not surprising, inasmuch as Russell's paradox involves some kind of self application and, as seen from the example of implication, proofs obviously are about themselves, specifically the proof p_0 is involved in some values of the variable p.

attached to which is the parenthesis we have seen:

(1.4c) (Incidentally, it is one of the peculiarities of constructive logic that, for some A, a *natural* formal proof of A goes *via* proofs of $A \rightarrow B$ and of $(A \rightarrow B) \rightarrow A$: such a proof of A actually contains a proof of $A \rightarrow B$.) (Kreisel & Newman, 1969, p. 58, original emphasis)[1]

With that, the paragraph ends.

[1] On p. 23 of the 'palimpsest of essays' Odifreddi (2019b), the passage (1.4a)–(1.4c) has retained its original content, but (1.4c) has been turned into a footnote.

These passages (1.4a)–(1.4c) raise, in order of reading, several questions, which will be discussed in one or more sections:[2]

1. What 'errors' are referred to in (1.4a)? (section 1.4.1)
2. In (1.4b) Kreisel notes that the informal construal of implication in (1.2) shares a dependence on 'some kind of self-application' with Russell's Paradox, but he does not go on to suggest that this would be reason to find (1.2) suspect. Whence the difference? (section 1.4.5)
3. What, in (1.4c), is the role of the explicit qualification of formality, what notion of naturalness is being appealed to, and what is meant by 'constructive logic'? (sections 1.2.1, 1.2.2, 1.2.3)
4. What are examples of such A and B as (1.4c) refers to? (sections 1.3.1, 1.3.3, 1.4.4)
5. Given that (1.4c) is about formal proofs, but (1.4b) about informal ones, what is the exact bearing of (1.4c) on (1.4b)? (sections 1.4.5, 1.5)
6. As far as I have been able to determine, Kreisel did not elaborate on (1.4a)–(1.4c) elsewhere. If that is correct, could a conjecture be made why he didn't? (section 1.4.5)

Possible differences between Kreisel's (1.2) and Heyting's (1.3) will be discussed as little as possible, as with respect to the remarks (1.4a)–(1.4c) and the questions they make one ask, they seem to be interchangeable.[3]

The main questions are those of the meaning of the terms involved (question 3) and of examples of such A and B (question 4). They determine the overall structure of this paper, as reflected in the table of contents.

1.2 Key concepts in that passage

1.2.1 Proof and formal proof

Kreisel characterises proofs, 'intuitive proofs' (Kreisel, 1968b, p. 321), as 'mental processes by means of which we convince ourselves of the validity of (mathematical) propositions' (Kreisel, 1973, p. 263), but also as objects that are 'abstract' (Kreisel,

[2] As sources for Kreisel's views that are pertinent to (1.4c), which was published in 1969, publications have been chosen from the encompassing period 1965–1973.

[3] The key terms describing differences are 'second clauses' and 'proof conditions versus assertion conditions'. The former refers to Kreisel's explicit demand for a mathematical proof (object) p_0; in Heyting's formulation this is either implicit or, on the contrary, absent because not required. Sundholm has argued for the latter, the idea being that what is required is rather an act (of understanding a construction process), which is not a mathematical object (1983, pp. 161–167, 169n13). To my mind, Sundholm is correct here. Be that as it may, it could be argued that (*i*) coming to possess a mathematical proof and coming to understand a construction process are both cases of coming to *accept*, that (*ii*) on either (otherwise unadorned) reading of Heyting's clause, accepting certain proofs of A presupposes accepting a constituent proof of $A \to B$, and that (*iii*) this captures enough of the 'involvement' that Kreisel speaks of in (1.4b).

1968b, p. 344) or 'intensional' (Kreisel, 1971a, p. 242). The senses of process and object are closely related, and not so much in opposition as revealing of the different ways in which proofs present themselves to us. In the order of both things and explanations, intuitive proof takes priority over formal proof:

(1.5) Indeed it is easy to forget that formal languages or formal derivations are introduced *because* they express propositions and proofs respectively: an argument which can be formalized by given derivation rules is conclusive *not* because the formalization proceeds according to *some* formal rules, but because the formal rules have been *seen* to preserve validity. Only in conjunction with this act of seeing the validity of the rules is the formal verification that a sequence of formulae is constructed according to given formal rules, a proof. In short, *proofs* as understood here (and in ordinary life and mathematics) are not linguistic objects. (Kreisel, 1970c, p. 29, original emphasis)

This was of course a staple in Brouwer, Heyting, and Gödel; among those of Kreisel's own generation, one finds a clear expression also in Myhill (1960), both authors – incidentally, born within five weeks of one another – having been strongly influenced by Gödel.[4]

Besides 'intuitive' and 'abstract', qualifications that are used (by Kreisel or others) to indicate proofs in this primordial sense are 'absolute', 'contentual', 'contentful', 'informal', or 'non-formal', depending on context and emphasis.[5]

[4] Of more recent defenses (or sympathetic discussions) of this idea, I here mention Leitgeb (2009), Tanswell (2015), and Crocco (2019).

[5] The latter four all serve to translate the German 'inhaltlich', which according to the *Etymologisches Wörterbuch des Deutschen* (https://www.dwds.de/d/woerterbuecher) goes back to the 17th century.

Further and later information about Kreisel's take comes from von Plato, who writes:

(1.6) I translate *inhaltlich* as contentful. Gödel suggested in the 1960s 'contentual,' but my translation is at least an English word. Georg Kreisel disliked it: He told me in July 2010 that one should just use the word *meaning*. *Inhaltlich*, then, would be *meaningfully*, or perhaps *in terms of meaning*. I regret not having asked what he thought of Gödel's invented word. (von Plato, 2017, p. 259, original italics)

In an email to me of November 7, 2020, von Plato adds that 'or perhaps *in terms of meaning*' is his adaptation for cases where 'meaningfully' would not fit, and that Kreisel preferred the translation to be as common a word in English as 'inhaltlich' is in German.

To take the claim about Gödel first, compare this remark by Van Heijenoort:

(1.7) A teaser for translators of German texts on foundations of mathematics is the word 'inhaltlich'. Mr. Bauer-Mengelberg coined the neologism 'contentual' and used it at a number of places. Elsewhere various periphrases were adopted; in particular, Professor Gödel suggested those that are used in the translation of his *1931*. (van Heijenoort, 1967, p. viii)

In that translation of Gödel's paper, it is always suitable locutions with 'meaning' or 'interpretation' that are used. As pointed out in Buldt (1997, p. 92), this indicates that Gödel did not like 'contentual', which word Van Heijenoort will surely have suggested to him, perhaps when they met in September 1963; the matter does not come up in their letters selected for Gödel (2003b). In the mentioned email, von Plato takes Buldt's point.

Two aspects of formalisation that come to mind as possible further determinants of the full meaning of 'formal proof' in (1.4c), individually or together, are explicitation and arithmetisation. While an informal reasoning is the starting point for a formal proof, the latter, when interpreted as intended, supplements it wherever needed by making explicit what was left implicit, such as certain premises or conditions. And through the device of arithmetisation, certain formulas whose intended meaning is arithmetical acquire additional meaning concerning properties of the formalism itself. (Such formulas may themselves be reasoned about informally or formally again.)

Below, I will use notation such as $A \to B$ both for meaningful propositions and for formal statements, and Gentzen's Natural Deduction to represent both informal and formal proofs. On each occasion, the context makes clear which is meant. The exact choice of proof system has no bearing on the present discussion, as long as it embodies the idea that if a formal proof is constructed on the basis of given ones, then the former retains the latter as parts; otherwise Kreisel's remark (1.4c) about 'literal containment' cannot apply. For example, Gentzen's Sequent Calculus (when seen as instructions for constructing Natural Deduction proofs) also complies, but tableau systems do not (Boolos (1984, pp. 377-378); D'Agostino and Mondadori (1994, p. 287)).

1.2.2 Naturalness

Although the adjective 'natural' has a variety of meanings, as witnessed in, for example, the *Oxford English Dictionary*, the one that seems the most suitable for a reading of (1.4c) is 'naturally arising or resulting from, fully consonant with, the circumstances of the case', as it is a naturalness that should reflect a 'peculiarity of

---

Although 'contentual' may well have seemed a neologism to all mentioned so far, strictly speaking it was not. The *Oxford English Dictionary* lists it with the meaning 'belonging to, or dealing with, content (opp. act or form)'. That is, admittedly, in its second *Supplement*, published in 1972 (and neither in the original volume for the letter C of 1893, nor in the first *Supplement* to the dictionary, of 1933), but the three citations given there are dated 1909, 1935, and 1962. (Incidentally, with an eye on Gödel's interests one notes that the citation of 1962, 'distinguishing the formal from the contentual features of propositions', comes from *Plato's Later Epistemology* (Runciman, 1962, p. 132) – where it is claimed that, for methodological reasons, Plato, unlike Aristotle, was not able to make that distinction.) In their translation of Weyl (1928), Bauer-Mengelberg and Føllesdal use 'contentual' for 'inhaltlich', but employ 'meaningful' for 'sinnerfüllt' (van Heijenoort, 1967, p. 482).

Curry (1941, p. 222) had translated 'inhaltlich' by his neologism 'contensive', chosen because it stands to 'content' as 'intensive' to 'intent'; the alternatives he mentions and rejects are 'material' and 'intuitive'. (Incidentally, Curry (1941, p. 224) considers contensive conceptions of mathematics (Platonism, intuitionism) to be 'useful only as secondary standards' compared to his primary standard of formalism.)

Finally, a reflection on Crocco's suggestion (2019) of a difference between Kreisel's 'informal proof' and Gödel's 'absolute proof' will have to wait for another occasion.

(I thank Jan von Plato for discussion of an earlier version of this footnote, and for permission to make use of his email.)

constructive logic'. Which among the 'circumstances' are the relevant ones, and what 'consonance' consists in, will depend on choices guided by values; these choices and values may moreover vary over time.[6] On the other hand, a formal proof can be completely described in factual or non-evaluative terms. A formal proof can therefore only be argued to be natural if it is seen in relation to something else, such that either that something or that relation is subject to evaluation, because an evaluative conclusion of an argument requires at least one evaluative premise. The obvious suggestion now is that if a formal proof derives its status of proof from an intuitive proof, then one way in which a certain formal proof could be considered natural is the derivative one where the intuitive proof it formalises is considered natural in some non-derivative sense. This is why, even though a discussion of natural constructive formal proofs (as in (1.4c)) is central to the present paper, its title leaves out the qualification 'formal'. What naturalness of an intuitive proof consists in is perhaps best considered case-by-case.

1.2.3 Constructivity

To have the idea of (*i*) formally constructive proofs to which (*ii*) constructive meanings are assigned according to certain explanations or interpretations, one must have notions of constructivity of two kinds, theoretical ones for (*ii*) and a pre-theoretical ('naive', 'heuristic') one for (*i*). The pre-theoretical notion is appealed to when characterising formalisms as constructive, and can be analysed into a variety of theoretical notions in terms of which these formalisms are then given full meanings; the appropriateness of each of the latter is subject to philosophical debate.[7] Of course, both the pre-theoretical and the theoretical notions have their uses also independently of any formalism.

Pre-theoretically, the following two familiar conditions on constructive non-formal proof are uncontentious: from a proof of an existential proposition one should be able to obtain an instance, and from a proof of a disjunction a proof of one of the disjuncts. Further conditions have led to debate. For example, in the development of intuitionistic logic, Johansson denied that Ex Falso holds for it, Freudenthal held that a proof of any proposition $A \rightarrow B$ must begin by proving A, and Griss argued that negation is not a constructive operation.[8] There is a large overlap with the concerns that led to the development of relevance logic, and it may be argued that Brouwer's

[6] A recent, detailed philosophical and linguistic analysis of naturalness in mathematics, culminating in an emphasis on 'the dynamic and prescriptive character of naturalness' that I find congenial, is Mauro and Venturi (2015). Note that they likewise consult a dictionary, 'the Oxford Dictionary' (presumably, given their quotations, the *Oxford Dictionary of English*), but only consider (in a careful way: see pp. 280 and 311) the primary definition they found there: 'Existing in or derived from nature; not made or caused by humankind', and not the fourth, 'in accordance with the nature of, or circumstances surrounding, someone or something'. The latter is closer to the meaning I appeal to.

[7] For an overview, see e.g. Ruitenburg (1991).

[8] For references and discussion, see van Atten (2017a).

ideas about logic (Brouwer, 1907, 1908) lead to a relevance logic (van Atten, 2009, p. 124). In the latter case, $A \to (B \to A)$ would not hold; certain instances may still be demonstrable, but not on the ground on which the schema is considered acceptable by others. (This will be relevant on p. 15.) Heyting (1956) accepted the debated principles mentioned. While Kreisel in his publications has little to say about this kind of discussion, his explicit point of reference for informal constructive logic is always Heyting.[9]

The statements of the informal conditions on existential and disjunctive propositions have formal analogues in what have become known as the Disjunction Property

$$\text{If } S \vdash A \vee B, \text{ then } S \vdash A \text{ or } S \vdash B\,. \tag{DP}$$

and the Existence Property

$$\text{If } S \vdash \exists x P(x), \text{ then } S \vdash P(t) \text{ for some term } t\,. \tag{EP}$$

The natural and common system HA has both (Kleene, 1945). But whereas the mentioned pre-theoretical conditions are constitutive of non-formal constructivism, it is not the case that, analogously, a formal system must have the properties (DP) and (EP) to count as formalisation of meaningful constructive thought. Here is Kreisel's proof for the case of (EP) (also for reference further on p. 15 below).

Theorem 1 *(Kreisel, 1970a, p. 125) There is a predicate P in the language of HA such that $\exists x P(x)$ is true on the intended interpretation, but the formal system $S = HA + \exists x P(x)$ does not have* (EP).

Proof We follow the proof in Troelstra (1973b, pp. 178-179). Define

$$R(x) = \overline{Prf}_{\mathrm{HA}}(x, \overline{\ulcorner \bot \urcorner}) \vee \forall y \neg \overline{Prf}_{\mathrm{HA}}(y, \overline{\ulcorner \bot \urcorner})\,, \tag{1.8}$$

where Prf_{HA} is a canonical proof predicate of HA.[10] From the consistency of HA we see that $\forall y \neg \overline{Prf}_{\mathrm{HA}}(y, \overline{\ulcorner \bot \urcorner})$ is true on the intended interpretation. Hence, with the decidability of $\overline{Prf}_{\mathrm{HA}}$, HA $\vdash \neg \overline{Prf}_{\mathrm{HA}}(\overline{n}, \overline{\ulcorner \bot \urcorner})$ for every $n \in \mathbb{N}$, and

$$\mathrm{HA} \vdash R(\overline{n}) \leftrightarrow \forall y \neg \overline{Prf}_{\mathrm{HA}}(y, \overline{\ulcorner \bot \urcorner}), \text{for every } n \in \mathbb{N}\,. \tag{1.9}$$

We also have, by the definition of R,

$$\mathrm{HA} \vdash \exists x R(x) \leftrightarrow (\exists y \overline{Prf}_{\mathrm{HA}}(y, \overline{\ulcorner \bot \urcorner}) \vee \forall y \neg \overline{Prf}_{\mathrm{HA}}(y, \overline{\ulcorner \bot \urcorner}))\,. \tag{1.10}$$

Consider the system $S = \mathrm{HA} + \exists x R(x)$. Trivially, $S \vdash \exists x R(x)$, and that formula is true on the intended interpretation, because, as noted, $\forall y \neg \overline{Prf}_{\mathrm{HA}}(y, \overline{\ulcorner \bot \urcorner})$ is; so

[9] On Freudenthal's conception, remark (1.4c) could not be made.

[10] $\ulcorner \psi \urcorner$ is the natural number that codes ψ in the given Gödel numbering, $\overline{n}$ is the representation of the natural number n in the formalism. $Prf_{\mathrm{HA}}(x, \ulcorner A \urcorner)$ holds if and only if x codes a proof of A in HA; $\overline{Prf}_{\mathrm{HA}}$ is its representation in the formal language of HA.

S, like HA, formalises meaningful constructive thought. Now suppose, towards a contradiction, that furthermore $S \vdash R(\overline{n})$, for some $n \in \mathbb{N}$. This is equivalent to

$$\mathrm{HA} \vdash \exists x R(x) \rightarrow R(\overline{n}) \tag{1.11}$$

for that n. Applying equivalence (1.10) to the antecedent and equivalence (1.9) to the consequent yields

$$\mathrm{HA} \vdash (\exists y \overline{Prf}_{\mathrm{HA}}(y, \overline{\ulcorner \perp \urcorner}) \vee \forall y \neg \overline{Prf}_{\mathrm{HA}}(y, \overline{\ulcorner \perp \urcorner})) \rightarrow \forall y \neg \overline{Prf}_{\mathrm{HA}}(y, \overline{\ulcorner \perp \urcorner}), \tag{1.12}$$

which simplifies to

$$\mathrm{HA} \vdash \exists y \overline{Prf}_{\mathrm{HA}}(y, \overline{\ulcorner \perp \urcorner}) \rightarrow \forall y \neg \overline{Prf}_{\mathrm{HA}}(y, \overline{\ulcorner \perp \urcorner}), \tag{1.13}$$

and then, using $\forall x \neg \varphi(x) \leftrightarrow \neg \exists \varphi(x)$ twice, to

$$\mathrm{HA} \vdash \forall y \neg \overline{Prf}_{\mathrm{HA}}(y, \overline{\ulcorner \perp \urcorner}). \tag{1.14}$$

But this contradicts the second incompleteness theorem. □

Complementarily, Troelstra pointed out that (EP) and (DP) are not sufficient conditions for a system to count as a formalisation of the intended constructive interpretation either: he gives the example of two extensions of HA, each of which has (EP), but which yield an inconsistent system when combined (Troelstra, 1973b, p. 179). (The extensions codify incompatible ideas about what constructive existence consists in.)

Given these results of Kreisel and Troelstra, I agree with Kreisel that the formal properties (DP) and (EP) 'are *not* linked to the constructive interpretation of Heyting's systems' (Kreisel, 1971c, p. 123, original emphasis).[11] Instead, I will take as criterion for the constructivity of a formal system: soundness with respect to the informal explanation of the meanings of the connectives in Heyting (1956).[12]

---

[11] I do not know whether Kreisel had found the results about (EP) and (DP) by the time of writing the Brouwer memoir, but the general point is made in Kreisel (1970a), which is the published version of his concluding address at the 1968 conference in Buffalo. In that same volume Prawitz presents a cut elimination theorem for a second-order system, points out the corollaries (DP) and (EP), and comments: 'These results, which certainly are consonant with intuitionistic principles, may have some bearing on the debated question whether a second-order system of the described kind is intuitionistic[ally] acceptable.' Then a footnote states: 'The intuitionistic significance of the described system has been advocated by Professor Kreisel especially. I am grateful to him for encouraging me to carry out these investigations' (Prawitz, 1970, p. 259). One possibility is that, by the time of the Buffalo conference, Kreisel had not yet developed his argument against (EP) and (DP) as criteria of 'intuitionistic significance'. A further possibility is that Kreisel's plea was first of all based on his view that quantification over species should be acceptable in ideological intuitionism; see section 1.4.5 below.

[12] This does not mean that one's intended interpretation must be Heyting's Proof Explanation, but it does rule out variations on it that validate formal systems for classical logic; on such variations, see Troelstra and van Dalen (1988, p. 9 and pp. 32–33, exercise 1.3.4) and Sundholm (2004). They depend on understanding basic notions such as construction and existence classically. Obviously Heyting, Brouwer, and Kolmogorov neither intended nor did this, and I therefore would rephrase

1.3 Cases depending on a proof of B

1.3.1 Finding the formula B from the formula A

In the search for an example of (1.4c), we may either start from ideas about such an A and from there attempt to reason towards the formulation of an appropriate B, or the other way around. In the first direction, this suggests the following formal proof skeleton, which is meant to represent not only, in the most straightforward way, the deductive relations indicated in (1.4c), but also, in the direction from left to right, the intended temporal order of construction of the two subproofs:

$$\dfrac{\dfrac{\begin{matrix}[A]^1\\ \vdots\\ B\end{matrix}}{A \to B}\,1 \qquad \dfrac{\dfrac{\dfrac{[A \to B]^2 \quad \dfrac{\begin{matrix}[B]^3\\ \vdots\\ A\end{matrix}}{B \to A}\,3}{A \leftrightarrow B} \qquad \begin{matrix}\vdots\\ B\end{matrix}}{A}}{(A \to B) \to A}\,2}{A} \tag{1.15}$$

Thus, we here have a formal proof of a given A that proceeds by a cut introduction with cut formula $A \to B$,[13] with the particularity that our choice for the formula B depends on our having first derived it from A. This leads to a second cut, with cut formula B.[14] If we now eliminate the first cut, we obtain

$$\dfrac{\dfrac{\dfrac{\begin{matrix}[A]^1\\ \vdots\\ B\end{matrix}}{A \to B}\,1 \quad \dfrac{\begin{matrix}[B]^3\\ \vdots\\ A\end{matrix}}{B \to A}\,3}{A \leftrightarrow B} \qquad \begin{matrix}\vdots\\ B\end{matrix}}{A} \tag{1.16}$$

This skeleton still brings out the simple fact that a sufficient condition for a formal proof of A to contain a subproof of $A \to B$, which in (1.4c) is the containment that is highlighted, is that A gets proved via proofs of $A \leftrightarrow B$ and B; this fact would

Troelstra and van Dalen's conclusion that 'This exercise shows that the BHK-interpretation in itself has no "explanatory power"', and also Sundholm's title 'The proof-explanation is logically neutral'. For these phrases to be correct, the 'BHK-interpretation' or 'proof-explanation' must be understood in a far more ambiguous (more widely schematic) sense than they had in their original contexts. As for the differences among the constructivists themselves (p. 6 above), note that the proposals of Johansson, Freudenthal, and Griss only lead to restrictions on Heyting's notion of construction, so that formal systems based on those proposals still satisfy the criterion I adopt.

[13] The term 'cut' comes from the Sequent Calculus, but I will use it also for the corresponding phenomenon in Natural Deduction (van Dalen, 2004, section 6.1).

[14] Usually, $\leftrightarrow$ is treated as a defined connective, and correspondingly a cut rule for it would be defined in terms of that for $\to$. For our present purpose, this is immaterial.

no longer be represented after elimination of the second cut. Although this second skeleton no longer fits (1.4c) literally, of course the essential part of the reasoning towards $(A \to B) \to A$ in the first skeleton is preserved.

With Curry's insight into the correspondence between implicational logic and the typing of functions, skeleton (1.15) can also be seen as a way to arrive at a judgement that an object a is of type A by finding a function $f\colon A \to B$, together with a selection functional $F\colon (A \to B) \to A$, where the latter is, in this case, obtained by finding a function $g\colon B \to A$ and an object b of type B. In that context, a motivation for transforming a proof in the other direction, from one with skeleton (1.16) into one with skeleton (1.15), would be to make the existence of that functional explicit in the formal proof.[15]

The question whether these formal proof skeletons may be natural will, as motivated in section 1.2.2 above, be approached by transposing the question to the informal. The informal counterpart to a cut introduction is a lemma introduction. A motivation for structuring an attempt at informally proving A as a proof of its equivalence to B together with a proof of B as a lemma would be that one expects or knows that, if one assumes that A is true, an equivalent B can be inferred that, if provable at all, should in some sense be easier to prove than A.[16] Two cases may be distinguished. The first is where this greater ease has its ground in the content of B, in case B is more explicit than A,[17] or is less complex,[18] or shifts to a more convenient domain;[19] for short: is better intelligible, and therefore, one expects, more fruitful to reason about.[20] Intelligibility is a value that depends on the content of the propositions involved in their relations to our capacity of understanding; it is a matter of degree, and furthermore for one and the same proof this degree may, with increased mathematical experience, change. The other case arises when B is among the stock of propositions already proved. There the value involved is efficiency of the simplest kind, independent of propositional content. Since in analysing (1.4c) we are after a 'peculiarity of constructive logic' as such, which has to do with (the contribution of the logic to the) propositional content, it is the first case that is of interest.

---

[15] For a comparative remark on the role of the selection functional here, see point 5 on p. 28 below. On selection functionals, see e.g. the introduction to Escardó and Oliva (2010).

[16] Dijkstra (1985, p. 7) writes that 'equivalence is the most underexploited connective in mathematics, in contrast to the implication that is used all over the place [. . .] the failure to exploit inherent symmetries often lengthens an argument by a factor of 2, 4 or more'. His explanation for this is that humans tend to reason in terms of cause and effect, and tend to assimilate implications to that. Note that to constructions in Brouwer's setting, the familiar distinction between tokens and types applies; so that if $A \to B$, at the token level there is a sense in which a construction for A causes the truth of B, but this is compatible with $A \leftrightarrow B$ at the type level. See also footnote 57 below.

[17] Through expansion of definitions, or by appeal to the meaning of A.

[18] See the remarks around (1.23) below.

[19] For example, by embedding $\mathbb{R}$ in $\mathbb{C}$. See also Kreisel (1967a, p. 166).

[20] Indeed, cut introduction is part of the proof restructuring that is done to make automatic, cut-free formal proofs, which are easier to find for computers, comprehensible to humans (e.g., D'Agostino et al. (2008)). On cut-free proofs, see section 1.3.2 below.

The heuristic of proving A by looking for such a (more) intelligible B is, not surprisingly, an old one, and goes back to the 'method of analysis and synthesis' of the Greek geometers.[21] The classical description was given by Pappus; well-known modern discussions outside the specialist literature[22] are Heath's in his edition of Euclid's *Elements* (Euclid, 1908) and Pólya's in *How to Solve It* (2004, first ed. 1945).[23]

Pappus writes:

> Now, analysis is the path from what one is seeking, as if it were established, by way of its consequences, to something that is established by synthesis. That is to say, in analysis we assume what is sought as if it has been achieved, and look for the thing from which it follows, and again what comes before that, until by regressing in this way we come upon some one of the things that are already known, or that occupy the rank of a first principle. We call this kind of method 'analysis', as if to say *anapalin lysis* (reduction backward). In synthesis, by reversal, we assume what was obtained last in the analysis to have been achieved already, and, setting now in natural order, as precedents, what before were following, and fitting them to each other, we attain the end of the construction of what was sought. This is what we call 'synthesis'. (1.17)
>
> There are two kinds of analysis: one of them seeks after truth, and is called 'theorematic'; while the other tries to find what was demanded, and is called 'problematic'. (Pappus of Alexandria, 1986, p. 82)

In bringing up Pappus, whose concern is with geometry, in the present paper, which is concerned with propositional and predicate logic, I do not mean to take a stand on the question whether the latter are the most appropriate vehicles for the representation

---

[21] Recent detailed linguistic analysis (Longa, 2020) has identified no fewer than 318 proofs in the Greek mathematical corpus (3rd century BC to 7th century CE) that apply the method of analysis and synthesis.

[22] As examples of that specialist literature in so far as it relates the method to modern constructive mathematics, I mention Knorr (1983), Mäenpää (1993), and Menn (2002). Brouwer also described this method in his dissertation; see p. 14 below.

[23] It is quite likely that Kreisel had seen those discussions. Kreisel reports that 'Since my school days I had had those interests in foundations that force themselves on beginners when they read Euclid's Elements (which was then still done at school in England)' (Kreisel 1989; trl. Odifreddi 2019a, p. 148). His school – Dudley (Isaacson, 2020, p. 90) – may well have used a different edition than Heath's, but if it was, then the latter would have been the first place to look further if one's interest had been piqued. The one reference to Pólya by Kreisel I have found is: 'The use of axiomatic analysis as a proof strategy does not seem to be well known to people writing on heuristics, like Polya' (Kreisel & Macintyre, 1982, p. 233). (Axiomatic analysis is a proof strategy because, once one has an axiomatic presentation of a subject, in trying to find a proof one need only take into account the properties mentioned in that presentation (Kreisel & Macintyre, 1982, p. 232).) That observation intimates a wide knowledge of Pólya's writings. Moreover they were colleagues at Stanford, where Kreisel taught from 1958-1959 and 1962-1985, incidentally the year Pólya died. Pólya had been an emeritus from 1953 on, but a very active one. Finally, perhaps Kreisel knew the remarks in Lakatos' series 'Proofs and refutations' (Lakatos, 1963–1964) (spread over part 1, pp. 10–11 note 2; part 3, p. 243n1, part 4, p. 305 and its footnote 1). In Kreisel's writings one finds references to the book of the same title that Lakatos published in 1976.

of the reasonings of the Greek geometers. Rather, the idea is that the method of analysis and synthesis is of wider applicability, and that various general remarks or phenomena that occur in the Greek context may also be pertinent or suggestive in domains where these logics are used.

From a modern constructivist point of view the difference between theorematic and problematic analysis is mostly one of perspective, as to find certain constructions for objects and their relations is what constructively proving the corresponding proposition consists in, and, conversely, proofs may themselves be regarded as mathematical objects one can look for.[24]

In its application to propositions, analysis is the process that leads from the assumption that A has been established to a proof of $A \to B$, where B has independently been recognised as provable, and synthesis the process of combining a proof of B and a proof of $B \to A$ to prove A. The heuristic turns on the fact that a consequence B of A corresponds to a necessary condition for the truth of A; if it is also a sufficient one, then, in a reversal of direction, what was construed as a proposition following from A is now construed as a proposition from which A follows. It is with an eye on this subsequent reversal that Pappus glosses 'analysis' as a 'reduction backward', and I take it that the latter sets things 'in natural order' because, by hypothesis, when embarking upon this reasoning B was already known, whereas A was not. The central idea in such a proof of A, then, is its equivalence to B. That B implies A is immediate if B is obtained from A by a chain of equivalences, but other cases will be more involved.[25] Be that as it may, the method will have fulfilled its heuristic function if one's proof of $A \to B$ at least suggests how to go about proving the converse.

The overall reasoning process will, in general, involve making guesses, but those will at least in part be motivated (directly or transitively) by the meaning of A.[26] As Pólya describes it, 'analysis is devising a plan, synthesis is carrying through the plan' (2004, p. 146), and this plan is visible in the successful outcome, if there is one. It also illustrates Rood's observation (made in a discussion of Kant) that 'if we look at proofs from a procedural point of view, then the boundary between discovery and justification starts to blur. A proof may itself involve various elements of discovery' (2005, p. 57). This makes proofs by analysis and synthesis highly intelligible and in that sense natural. When the reasoning steps employed in a successful use of this heuristic method can be mirrored and explicated in formal inferences in a suitable system, the skeleton of a formal proof then is that as those above.

Perhaps one takes an alternative representation of the informal reasoning by analysis and synthesis to be given by the *pair* of proofs

[24] As set out in Kolmogorov (1932), which, as he succinctly put it later, 'was intended to construct a unified logical apparatus dealing with objects of two types – propositions and problems.' (2019, p. 452); see also Bernays' review of Pólya (Bernays, 1947, 184–185).

[25] A discussion of this latter point in the Greek context is Behboud (1994).

[26] A recent study of 'motivated proofs' is Morris (2020).

$$\dfrac{\begin{matrix}[A]^1 \\ \vdots \\ B\end{matrix}}{A \to B}\,1 \qquad\qquad \dfrac{\dfrac{\begin{matrix}[B]^1 \\ \vdots \\ A\end{matrix}}{B \to A}\,1 \qquad \begin{matrix}\vdots \\ B\end{matrix}}{A} \tag{1.18}$$

in which, crucially, the proof of $A \to B$ does not figure as a subproof in the proof of A. Indeed, in the corpus of ancient Greek geometry one finds proofs where only an analysis is presented, or only a synthesis. The analysis sufficed in case the subsequent synthesis was considered to be obvious,[27] the synthesis if the only interest was in a deductive justification of A.[28] But it was realised that giving both facilitates understanding, and hence can make proofs more convincing. In addition, there is a rhetorical value to such a double presentation: it '[creates] the illusion that the solution is necessary and *emerges naturally* out of the problem'. (See for these two points Netz (2000); the emphasis here is mine.) Evidently a presentation of neither analysis nor synthesis alone can count as a full rendering of the kind of reasoning process under discussion, and even the pairwise formalisation above remains incomplete, as it distributes the representation of the reasoning towards one conclusion over two proofs. Either stands on its own, and the intended relation between those two, in (1.18) suggested by the juxtaposition and our choice of the same schematic letters, is not itself represented formally. In particular, the central idea of a proof by analysis and synthesis, the equivalence of A to B, is absent. Combining both trees into one, as in skeleton (1.16), makes it appear, and thereby Kreisel's claimed containment relation. That containment is not an artefact of the representation, but results from an explicitation of the top-level structure of the informal, intelligible reasoning that it represents.

A whole class of examples of proofs with skeleton (1.16) is generated by a formal theory for which quantifier elimination has been established constructively, and for which the quantifier-free statements obtained are decidable. Then an appropriate B can be found from A without any need for guessing: one has an algorithm that for every sentence A yields a quantifier-free sentence B such that $A \leftrightarrow B$ is provable in the theory and B is decidable.[29] A positive outcome of the corresponding decision

---

[27] Or, in the case of reductio ad absurdum, superfluous (this is implicit in Pappus' account, see Mäenpää (1993, 188n83)).

[28] See Knorr (1986, p. 9) for the general point, and its p. 377n89 for some examples. Incidentally, Knorr received support from Kreisel:

> Knorr read the ancient texts so as to reveal new proof strategies; indeed he dated different bits even of Euclid in terms of a more or less linear development of proof ideas, which of course does not conform to the present order of the Elements. Oral history: the proof theorist Georg Kreisel, Knorr's colleague at Stanford [Knorr joined the faculty in 1979], strongly encouraged this work. I believe he did so because it pointed not only to the history of proof but also to the human discovery of capacities for proving. (Hacking, 2014, p. 140) (1.19)

[29] E.g., (classical and constructive) Presburger Arithmetic (Presburger 1930; Kleene 1952, pp. 204, 474–475), or the classical theory of real closed fields (Lombard & Vesley, 1998; Tarski, 1948).

procedure for B then leads to a proof of A.[30] It need of course not be the case that the quantifier-free statement is, as a whole, better intelligible to humans than the quantified equivalent, quite the contrary.

Of course, quantifier elimination and also the analytic-synthetic method exist in a classical context as much as in a constructive one, and as such cannot count as an interpretation of (1.4c), which asks for a 'peculiarity of constructive logic'. But in the case of the analytic-synthetic method, one can be found by narrowing down the scope.

Consider the characteristic demand of constructive logic that existential statements $\exists x P(x)$ be proved by exhibiting an a such that $P(a)$. A well known heuristic to fulfill that demand can readily seen as an application of the method: assume that $\exists x P(x)$ and then attempt to reason, from the definition of the predicate P and other available information (axioms and previously obtained theorems) towards conditions that a witness for it must satisfy, in such a way that combining them leads to conditions that are moreover sufficient. Let a be a hypothetical object satisfying those conditions; thus one has shown that, if the existential statement has a witness at all, then a must be one.[31] The second part of the attempt consists in establishing that an a fulfilling these conditions can indeed be constructed. This is how Brouwer described it in his dissertation:

(1.20) There is a special case [. . .] which really seems to presuppose the hypothetical judgment from logic. This occurs where a building in a building is defined by some relation, without that relation being immediately seen as a means for constructing it. Here one seems to *assume* to have effected the construction looked for, and to deduce from this hypothesis a chain of hypothetical judgments. But this is no more than apparent; what one is really doing here, consists in the following: one starts by constructing a system that fulfills part of the required relations, and tries to deduce from these relations, by means of tautologies, other relations, in such a way that in the end the deduced relations, combined with those that have not yet been used, yield a system of conditions, suitable as a starting-point for the construction of the system looked for. Only by that construction will it then have been proved that the original conditions can indeed be satisfied. (Brouwer, 1975, p. 72, emphasis Brouwer, translation modified)[32]

---

[30] Carrying out decision procedures based on quantifier elimination may quickly become unfeasible for humans and then also for computers; here our concern is with the principle.

[31] This is reminiscent of the classically provable 'dual version of The Drinking Principle': 'There is at least one person such that if anybody drinks, then he does' (Smullyan (1978, p. 210); Warren et al. (2018)). But here we are reasoning constructively and, in particular, without appeal to Ex Falso, and conclude something only about a hypothetical object.

[32] Er is een bijzonder geval [dat . . .] werkelijk het hypothetische oordeel der logica schijnt te vooronderstellen. Dat is, waar een gebouw in een gebouw door eenige relatie wordt gedefinieerd zonder dat men daarin direct het middel ziet het te construeeren. Het schijnt, dat men daar *onderstelt* dat het gezochte geconstrueerd was, en uit die onderstellingen een keten van hypothetische oordeelen afleidt. Maar meer dan schijn is dit niet; wat men hier eigenlijk doet, bestaat in het volgende: men begint met een systeem te construeeren, dat aan een deel der geëischte relaties voldoet, en tracht

Brouwer's main interest here is in the question whether in a mathematics founded on intuitive givenness, hypothetical constructions can be made sense of at all. His answer is positive: hypothetical statements should be construed not as propositional expressions of possible, as yet unknown truths, but as conditions (on constructions). Conditions that we pose are themselves actual objects, and no bearers of truth or falsehood.

For the present purpose, the interest is elsewhere, namely in the structure of the reasoning that Brouwer describes. In a footnote, he give as examples 'the uniqueness proofs for transformation groups with given properties by Hilbert and Lie, and also ordinary elementary problems, such as looking for a common harmonic pair, or the problems of Apollonius'; the latter being classical examples of the application of the method of analysis and synthesis.[33] (Often, the existence of the 'building' in which another 'building' is to be constructed is itself given as an hypothesis, and specified in terms of parameters; then what needs to be proved rather takes a form like $R(n) \to \exists x P(x, n)$. See below, at the discussion of (1.24).) Now, as before: Suppose that the reasoning employed in a successful case can fully be mirrored in formal inferences in a given system. Then the candidate witness a can be obtained from the hypothesis using the inferential resources of the system itself, so that this formal proof of $\exists x P(x)$ begins with a proof of $\exists x P(x) \to P(a)$. Taking $\exists x P(x)$ for A and $P(a)$ for B, we see that the skeleton of the overall proof is that in (1.16).

Where this works out, the formal system is able to reflect the relation between the propositions $\exists x P(x)$ and $P(a)$ to a greater extent than applications of (EP) are able to. We refer back to Kreisel's proof that (EP) is not a necessary condition for the constructivity of a formal system (p. 7). That proof also shows that closure of a formal system under the rule (EP) does not guarantee provability in the system of the corresponding implication: On the one hand, (EP) is a schematic conditional that is constructively correct also when instantiated with the hypothesis $\mathrm{HA} \vdash \exists x R(x)$; this is not changed by the fact that we know that, because of the consistency of HA, that hypothesis will never be fulfilled. On the other hand, again because of the consistency of HA, we know that for no t the corresponding implication can be proved in it.

And even in cases where we have a system with (EP) and $\vdash \exists x P(x)$, for some P, a particular proof in that system of $\exists x P(x) \to P(t)$, for some t, need not proceed by explicating a connection between its antecedent and its consequent. HA has (EP) and contains the axiom $A \to (B \to A)$. Suppose that we have a P such that $\vdash \exists x P(x)$. One obtains $\vdash P(t)$ for some t by (EP), and then, via an instantiation of the axiom with $A = P(t)$ and $B = \exists x P(x)$, $\vdash \exists x P(x) \to P(t)$. But the axiom depends on no relation between A and B at all (which for relevantists is the reason to reject it). It may be observed that nevertheless a relevant connection between $\exists x P(x)$ and t has

uit die relaties door tautologieén andere af te leiden zóó, dat tenslotte de afgeleide zich met de nog achteraf gehoudene laten combineeren tot een stelsel voorwaarden, dat als uitgangspunt voor de constructie van het gezochte systeem kan dienen. Met die constructie is dan eerst bewezen, dat werkelijk aan de voorwaarden kan worden voldaan. (Brouwer, 1907, pp. 126–127)

[33] Incidentally, Brouwer's description gives the impression that he had seen the discussion of analysis and synthesis in Hankel (1874, esp. p. 141).

been exploited, namely the one that is made in the proof, at the metalevel, of (EP) itself. However, the point is that, while the construction method for the formal proof of $\exists x P(x) \to P(t)$ described here depends on that connection, the formal proof obtained does not represent it.

The relation between constructive existential quantification and the method of analysis and synthesis naturally extends to choice principles. Given the informal constructive meaning of the quantifiers,

$$\forall x \in D\, \exists y \in D' P(x, y) \tag{1.21}$$

must be proved by providing a method that transforms any proof of $a \in D$ into a b together with proofs of $b \in D'$ and $P(a, b)$.[34] Thus, if such a method is embodied in a constructive function $f : D \to D'$, a natural way of establishing (1.21) consists in observing that

$$\forall x \in D\, \exists y \in D' P(x, y) \leftrightarrow \exists f\, \forall x P(x, f(x)) \tag{1.22}$$

and then proving the right hand side. If this reasoning can be fully formalised, the result is a natural formal proof of (1.21) with skeleton (1.16), taking for A and B the left and right hand sides of (1.22).

This naturalness would explain Kreisel's observation, in correspondence with Mints, that various propositions seem easier to prove when understood constructively than when understood classically. That observation is reported in Kosheleva and Kreinovich (2015), where an explanation is offered that is based on the fact that constructively (and under the Church-Turing Thesis), Π^0_2-statements about the natural numbers are equivalent to Σ^0_2-statements

$$\forall x \in \mathbb{N}\, \exists y \in \mathbb{N} P(x, y) \leftrightarrow \exists e \in \mathbb{N}\, \forall x \in \mathbb{N} P(x, \{e\}(x)), \tag{1.23}$$

where e is the index of a recursive function and $\{e\}(x)$ denotes the result of its application to x. On this reading, the original statement is therefore in the class of Δ^0_2-statements, and should be expected to be easier to prove than on its classical classical understanding.[35] This is a special case of that in the preceding paragraph.

Informal existence theorems that depend on a hypothesis are typically of the form

$$\forall x \in D(R(x) \to \exists y \in D' P(x, y)). \tag{1.24}$$

On predicate-logical reconstructions, solved 'problems' in ancient Greek geometry are examples; on this point, see Menn (2002, pp. 202–204). A constructive proof of such a theorem calls for a method that, dependent on proofs of $n \in D$ and $R(n)$, constructs a proof of $\exists y \in D' P(n, y)$. To find such a method, one treats these

[34] Pitfalls in justifying such a choice principle for other domains than the natural numbers are discussed in Troelstra and van Dalen (1988, pp. 189-190).

[35] Kreisel will also have had in mind examples of a different kind, in the theory of recursive ordinals, and depending on specifically intuitionistic notions: see Kreisel (1965, p. 2.6141) and Kreisel (1983, pp. 228–229, not in the 1958 version).

dependencies as additional information in a parameter n, and may follow the (first part of the) strategy for proving existential statements without an hypothesis sketched above. If this works out, one obtains an a such that

$$\forall x \in D(R(x) \to (\exists y \in D' P(n, y) \leftrightarrow (a \in D' \wedge P(n, a)))) \,. \tag{1.25}$$

If this method does not actually depend on the way in which $n \in D$ and $R(n)$ have been proved, it furthermore serves to establish

$$\exists f \colon D \to D' \, \forall x \in D(R(x) \to P(x, f(x))) \,, \tag{1.26}$$

which implies, and is implied by,

$$\forall x \in D \, \exists y \in D'(R(x) \to P(x, y)) \,. \tag{1.27}$$

This leads to a natural formal proof of (1.27) using (1.26) and skeleton (1.16).

1.3.2 Digression: a Kreiselian objection

In its application to formal proofs, the term ‘analytic’ nowadays is also used with a different meaning, assigned to it in Smullyan (1965). The difference was remarked on by Kreisel as soon as it arose. In *Mathematical Reviews*, he wrote:

> The author introduces the very happy terminology of ‘analytic’ deduction, i.e., deduction involving the analysis (breaking up) of assertions. This replaces the less elegant ‘deduction rules possessing the subformula property’, or the quite misleading ‘cut free’ (misleading because it refers to a specific formalization). The terminology seems good despite its conflict with the traditional use of ‘analytic’ in contrast to ‘synthetic’ reasoning. (Kreisel, n.d.-d) (1.28)

The conflict consists in the fact that formal proofs that are analytic in this new sense of the term never prove A via a subproof of $A \to B$, whereas the presence of the latter is characteristic for formalised Greek analysis as discussed in section 1.3.1.

As the conflict is terminological, it has no bearing on the argument there to the effect that certain formal proofs with that pattern are natural because the meaningful proofs by analysis and synthesis that they formalise are. However, for some years Kreisel made a claim about the relation between analytic proofs in the new sense and meaningful proofs that would have a bearing on that argument. It is discussed here in the form of a digression, mainly because Kreisel came to see that the claim was based on a mistake.

The context of his claim is formal systems for which it can be demonstrated that all their proofs can be transformed (‘normalised’) into ones that are analytic in the new sense; a corollary of such a demonstration is that these systems have the properties (DP) and (EP) (p. 7 above). In 1967, Kreisel writes, somewhat tentatively:

(1.29) [C]onsider now the rule:

With every formal derivation D in F of an existential (numerical) formula, i.e. a closed formula of the form $\exists x A(x)$, associate *that* x which is supplied by the proof *which you understand to be represented by* D (x will, in general, be a term containing parameters).

[...]

Reflection shows that when one thinks through a formal argument D in Heyting's system, the *thought* involved is more closely represented by the cut free proof D' associated by means of so-called cut elimination [8] with D: D' has the property that if it proves $(\exists x)A(x)$, it *mentions* a particular $A(t)$ from which it obtains directly $(\exists x)A(x)$. Thus, though D' may still be understood differently by different people, it is a detailed enough representation of the intuitive thought to settle the particular question above (namely, what x is supplied by the proof that we understand to be represented by D?). All that is needed is this: each of us should convince himself that cut elimination provides a correct (i.e. more faithful) analysis of the proof which D represents for us.[36] (Kreisel, 1967b, pp. 244-245, italics his)

By 1971, in his review of Gentzen's *Collected Papers*, he is fully confident:

(1.30) To every derivation d there is a normal derivation $|d|$ that expresses the same proof as d. (Thus 'normal' derivations provide canonical representations, roughly as the numerals provide canonical representations for the natural numbers.) And natural-deduction systems are distinguished at least to the following extent: the particular normalization steps for which we get normal forms that are independent of order, *evidently* preserve the proof expressed by the derivation (to which the step is applied). (Kreisel, 1971a, p. 245, original italics)

And, in a publication of the same year,

(1.31) A minimum requirement is then that *any derivation can be normalized*, that is transformed into a unique normal form by a series of steps, so-called 'conversions', each of which preserves the proof described by the derivation. This requirement has a formal and an informal part:

(α) The *formal* problem of establishing that the conversions terminate in a unique normal form (independent of the order in which they are applied).

(βi) The *informal* recognition (by inspection) that the conversion steps considered preserve identity, and the informal problem of showing that

(βii) distinct, that is incongruent normal derivations represent different proofs (in order to have unique, canonical, representations).

[36] The current view of cut elimination is technical; the view I propose above is probably not shared generally. But it should be noted that though it is more precise and specific (and, perhaps, wrong in detail), it is not inconsistent, for instance, with Brouwer's view, [1], footnote 8, where he speaks of fully analysed, canonical proofs and stresses that they are infinite structures. [Note MvA: The reference '[1]' is to 'Über Definitionsbereiche von Funktionen' (Brouwer, 1927a).]

> For examples of remarkable progress with the formal problem see the work of Martin-Löf and Prawitz in this volume.[37] The particular conversion procedures considered evidently satisfy requirement (βi) since each conversion step merely contracts the introduction of a logical symbol immediately followed by its elimination. Such a contraction clearly does not change the proof described by the two formal derivations (before and after contraction).
>
> [. . .]
>
> As stressed by Prawitz [1971], his normalization procedures obviously preserve identity of proofs.[38]
>
> (Kreisel, 1971c, pp. 112, 114–115, original emphasis)

On the view Kreisel expresses here, two formal proofs that differ in that one is the normalisation of the other still represent the same (understood) abstract proof. Now, by Kreisel's earlier claim quoted in (1.4c) above, there are cases, depending on A, where a non-normalised formal proof can be said to be the natural one. While these two ideas are not necessarily in conflict – there may, for a given notion of naturalness, be both natural and unnatural representations of one and the same object –, they are so on the strongly procedural view on proofs taken above, according to which formal proofs are (primarily) construed as representations of acts of reasoning: applying normalisation rules to a formal proof that reflects an intuitive proof obtained by acts of analysis and synthesis makes that reflection vanish.[39] From this perspective, one would, in interpreting the method of analysis and synthesis, not exploit the subformula property of certain formal systems,[40] and whether a proof of $\exists x A(x)$ proceeds by presenting an instance would be too crude a criterion for identifying the thought involved in that proof, contrary to Kreisel's proposal in (1.29) above.[41] This could of course be turned around, and then Kreisel's claim would be seen as an objection to the strongly procedural view on proofs taken here.

However, Kreisel soon came to reject the view expressed in (1.29)–(1.31). In Kreisel and Takeuti (1974, p. 38n8), with reference to the Gentzen review as quoted above in (1.30), he calls it 'even more implausible than appears from [that review]', precisely because of the distance of normalised proofs from '[t]he faithful representation of actual reasoning' (Kreisel & Takeuti, 1974, pp. 37, 38n8). And in an unpublished postscript to that review, he calls the claim 'evidently false', accepting Statman's diagnosis that he had confused the preservation of identity, i.e.,

$$\underline{d}_1 = \underline{d}_2 \Rightarrow \underline{|d_1|} = \underline{|d_2|} \tag{1.32}$$

---

[37] [footnote MvA] Fenstad (1971).

[38] [footnote MvA] In fact Prawitz' attitude in the place Kreisel refers to is more like Troelstra's: he speaks of a 'conjecture', 'a reasonable thesis' (1971, p. 257).

[39] In more recent work on the question of the identity of proofs, the presence of lemmas is considered to be a distinguishing feature; see e.g. Straßburger (2019, section 2(a)).

[40] Hintikka and Remes (1976) do this; for criticism on this point also e.g. Behboud (1994, p. 61).

[41] Interesting middle ground here may be provided by the so-called analytic cut rule, in which the cut formula has to be a subformula of the assumptions or the conclusion (D'Agostino & Mondadori, 1994; Smullyan, 1968).

(if the abstract proofs expressed by the derivations d_1 and d_2 are the same, then so are the abstract proofs expressed by the respective normalisations) and

$$\underline{d} = \underline{|d|} \tag{1.33}$$

(the abstract proof expressed by a derivation d is the same as the abstract proof expressed by the normalisation of d) (Kreisel, 1976, pp. 6–7).

1.3.3 Finding the formula *A* from the formula *B*

There are also cases where an A is found starting from considerations about a certain B (or, initially, an open $B(x)$). A formal context that provides examples here, depending on arithmetisation, is the Diagonal Lemma or Fixed Point Lemma (for arithmetic). Consider the following standard formulation and proof:

Theorem 2 (Diagonal Lemma for Formulas) *Let S be a system that contains primitive recursive arithmetic. Then for each formula $\varphi(x)$ with only x free there exists a sentence ψ such that $S \vdash \psi \leftrightarrow \varphi(\overline{\ulcorner \psi \urcorner})$.*

Proof (after van Dalen (2004, p. 251)). Let $s(x, y)$ be a primitive recursive function such that $s(\ulcorner \varphi \urcorner, \ulcorner t \urcorner) = \ulcorner \varphi[t/x] \urcorner$, so that s is a substitution function specialised to substitutions for the variable x. Let the predicate $\sigma(x, y, z)$ represent $s(x, y)$ in the formalism. Define $\theta(x) = \exists y(\varphi(y) \wedge \sigma(x, x, y))$, $m = \ulcorner \theta(x) \urcorner$, and $\psi = \theta(\overline{m})$. The definitions give the immediate equivalences

$$\vdash \psi \leftrightarrow \theta(\overline{m}) \,, \tag{1.34}$$

$$\vdash \theta(\overline{m}) \leftrightarrow \exists y(\varphi(y) \wedge \sigma(\overline{m}, \overline{m}, y)) \,, \tag{1.35}$$

and

$$\vdash \psi \leftrightarrow \exists y(\varphi(y) \wedge \sigma(\overline{m}, \overline{m}, y)) \,. \tag{1.36}$$

As σ represents s,

$$\vdash \forall y(\sigma(\overline{m}, \overline{m}, y) \leftrightarrow y = \overline{s(m, m)}) \,, \tag{1.37}$$

and by the definition of m

$$\vdash \forall y(\sigma(\overline{m}, \overline{m}, y) \leftrightarrow y = \overline{\ulcorner \theta(\overline{m}) \urcorner}) \,. \tag{1.38}$$

With this we obtain, from (1.36),

$$\vdash \psi \leftrightarrow \exists y(\varphi(y) \wedge y = \overline{\ulcorner \theta(\overline{m}) \urcorner}) \,, \tag{1.39}$$

hence

$$\vdash \psi \leftrightarrow \varphi(\overline{\ulcorner \theta(\overline{m}) \urcorner})\,, \tag{1.40}$$

and then by the definition of ψ

$$\vdash \psi \leftrightarrow \varphi(\overline{\ulcorner \psi \urcorner})\,. \tag{1.41}$$

□

For the present purpose, I modify this proof slightly, and from (1.38) first obtain, by predicate logic,

$$\vdash \exists y(\varphi(y) \land \sigma(\overline{m}, \overline{m}, y)) \leftrightarrow \varphi(\overline{\ulcorner \theta(\overline{m}) \urcorner}) \land \sigma(\overline{m}, \overline{m}, \overline{\ulcorner \theta(\overline{m}) \urcorner})\,, \tag{1.42}$$

which, since by (1.38)

$$\vdash \sigma(\overline{m}, \overline{m}, \overline{\ulcorner \theta(\overline{m}) \urcorner})\,, \tag{1.43}$$

then reduces to

$$\vdash \exists y(\varphi(y) \land \sigma(\overline{m}, \overline{m}, y)) \leftrightarrow \varphi(\overline{\ulcorner \theta(\overline{m}) \urcorner})\,. \tag{1.44}$$

Now with (1.36) we find ourselves at (1.40) again. The point of this detour is that the proof now passes, in (1.42), through a proof of the equivalence of a certain existential statement (namely, ψ) to one of its instances; so that ψ is an A as in (1.4c), taking the right-hand side of (1.42) for B (this illustrates the reflection on (1.20) above).

While the proof could be simplified if the language contains a function or term for s instead if representing it by σ, it is the representation that allows for the introduction of the existential quantifier that my point depends on. Representation brings out the existential quantification in the notion of functionality. (For a proof simplified in this sense, see the Diagonal Lemma for Terms and its corollary in section 1.4.2.)

When applying the Diagonal Lemma, one reasons from or toward ψ via $\varphi(\overline{\ulcorner \psi \urcorner})$ in one step, leaving the passage through the instance of ψ on which that step depends implicit (if the theorem is proved in the modified way); and one may have proved it differently. Be that as it may, if the goal of the application is to establish ψ itself, then as long as $\varphi(\overline{\ulcorner \psi \urcorner})$ is proved from yet another equivalent, the overall proof of ψ this yields retains the form of skeleton (1.16). For example, consider a proof of Gödel's incompleteness theorem for a (consistent) system S via an application of the Diagonal Lemma to $\neg\, Pr_S(x)$. The fixed point ψ this yields is undecidable in S, but provable in a suitable system U in a proof with skeleton (1.16), where A is ψ and B is, for example, $Con(S)$ or the reflection principle $Pr_S(\overline{\ulcorner \gamma \urcorner}) \to \gamma$, for closed $\gamma \in \Pi^0_1$. That principle is studied in Kreisel's joint paper with Levy (1968). They are convinced that

> What makes reflection principles useful is that they have a clear intuitive meaning, and so, if such a principle is provable in U, we have a good chance of finding a proof. (1.45)

This, in effect, connects reflection principles and the considerations on intelligibility on p. 10 above.[42]

A related application of the Diagonal Lemma is

Theorem 3 (Löb) *(Löb, 1955) Let S be a system that contains recursive arithmetic and a provability predicate Pr_S satisfying certain natural conditions.*[43] *If $S \vdash \overline{Pr_S}(\overline{\ulcorner\varphi\urcorner}) \to \varphi$, then $S \vdash \varphi$.*

Where, for given φ, Löb in his proof had appealed to the Diagonal Lemma to obtain a fixed point ψ such that

$$\vdash \psi \leftrightarrow (\overline{Pr_S}(\overline{\ulcorner\psi\urcorner}) \to \varphi)\,, \tag{1.46}$$

Kreisel devised a variant proof using a fixed point ψ such that

$$\vdash \psi \leftrightarrow \overline{Pr_S}(\overline{\ulcorner\psi \to \varphi\urcorner})\,. \tag{1.47}$$

This equivalence does not mean that a natural formal proof of ψ contains a proof of $\psi \to \varphi$, because of the indirection introduced by the (formal) provability predicate, which refers to (formal) proofs only through a coding. Similarly, if one first tries to prove $\psi \to \varphi$ and then appeal to

$$S \vdash \overline{Pr_S}(\overline{\ulcorner\gamma\urcorner}) \text{ exactly if } S \vdash \gamma \tag{1.48}$$

to arrive at a proof of ψ, the relation holding between the formal proofs obtained is not that of containment. This could not be changed by adding an axiom schema $S \vdash \overline{Pr_S}(\overline{\ulcorner\gamma\urcorner}) \leftrightarrow \gamma$ to the system, because that is inconsistent, and would not be changed by adding the admissible rules corresponding to (1.48), because the device of an admissible rule introduces indirection in its own way: it indirectly presents a proof that uses only the rules that are constitutive of the system. However, Kreisel also used his fixed point to prove

Theorem 4 (Formalised Löb) *(Kreisel & Takeuti, 1974, pp. 44–45) Let S be a system as required for Löb's Theorem. Then $S \vdash \overline{Pr_S}(\overline{\ulcorner\overline{Pr_S}(\overline{\ulcorner\varphi\urcorner}) \to \varphi\urcorner}) \leftrightarrow \overline{Pr_S}(\overline{\ulcorner\varphi\urcorner})$*

Being an equivalence, this is a strengthening of what Formalised Löb's Theorem would strictly be. If one now attempts to obtain a formal proof of $Pr_S(\ulcorner Pr_S(\overline{\ulcorner\varphi\urcorner}) \to \varphi\urcorner)$ it would, since the right hand side is conceptually simpler than the left hand side, be natural to do so via a formal proof of $Pr_S(\ulcorner\varphi\urcorner)$ (and to obtain the latter, if $\varphi \in \Sigma^0_1$, by proving φ and appealing to Σ^0_1-completeness of S, which is a condition for

---

[42] By a coincidence, the final manuscript of their paper was accepted within days of Brouwer's death (December 8 and December 2, respectively; see Kreisel and Lévy (1968, p. 142)), so that Kreisel had such applications well in mind when setting out to write the obituary from which (1.4c) is taken. A letter from Kreisel to Heyting of February 20, 1967 shows that by then preparation for Kreisel and Newman (1969) had begun; and a postcard in the same direction of April 14, 1969, that by then Kreisel was still working on it (Heyting Papers, Bkre 690220, Bkre 690414).

[43] See, besides Löb (1955) and Kreisel and Takeuti (1974), e.g. Smoryński (1991).

Kreisel's proof). Thus, $Pr_S(\ulcorner Pr_S(\overline{\ulcorner\varphi\urcorner}) \to \varphi\urcorner)$ again is an A exemplifying (1.4c), found from $B = Pr_S(\ulcorner\varphi\urcorner)$.[44]

1.4 Cases independent of a proof of *B*: proof paradoxes

1.4.1 'Errors'

As we saw in the introduction, the immediate context in which Kreisel makes his remark (1.4c) is one in which he states that implication, when understood as an operation on contentual proofs and not (only) a formal connective, invites 'errors which are, formally, similar to Russell's paradox in set theory'. He does not specify any, but, by the time of writing the Brouwer obituary, he knew, through his own work and personal contacts:

1. his paradox in an untyped λ-calculus enriched with 'notions' (section 1.4.4.1);
2. Gödel's Paradox in Church's system of 1932-1933 (intuitionistic version) (section 1.4.4.2);
3. Troelstra's Paradox in the 'theory of the Creating Subject' (section 1.4.4.3);
4. Goodman's Paradox in the 'theory of constructions' (section 1.4.4.4).

References are given in the dedicated subsections below. The main interest will be in how features they have in common make them illustrate Kreisel's remarks (1.4a)–(1.4c):

1. Each of these paradoxes turns on the existence of a particular proof whose existence is, in effect, concluded to by an application of Lawvere's Fixed Point Theorem.
2. In each the existence of that particular proof entails

$$A \leftrightarrow (A \to \perp) \tag{1.49}$$

   for a certain proposition A.

The (extent of the) formal resemblance to Russell's Paradox is clear if in the formulation of these two features one takes sets instead of proofs for the objects. The first feature distinguishes these *proof* paradoxes from the *provability* paradoxes, e.g., 'This proposition is not provable' or the Myhill-Montague Paradox; I will briefly return to these at the end of this section.[45]

---

[44] It should be noted that the proof of the equivalence depends, in the direction from right to left, on the acceptance of $A \to (B \to A)$. As recalled on p. 7 above, this is not acceptable on every view of constructive logic.

[45] In their discussion of Goodman's Paradox, Dean and Kurokawa (2016), following a suggestion of Weinstein (1983, p. 264), emphasise the (as they are aware, limited) extent to which it resembles Montague's; the present approach, the extent to which it is different.

It can be argued that in the reasonings embodied in these proof paradoxes an error has been made, to the extent that it can be argued that (1.49) leads to $\perp$. It is a well known general fact that if one has A and B such that

$$A \leftrightarrow (A \rightarrow B)\,, \tag{1.50}$$

then positive implicational logic suffices to prove A and then B from the two component implications

$$A \rightarrow (A \rightarrow B) \tag{1.51}$$

and

$$(A \rightarrow B) \rightarrow A\,. \tag{1.52}$$

One first derives the contraction of (1.51) by

$$\dfrac{\dfrac{[A]^1 \qquad \dfrac{[A]^1 \qquad A \rightarrow (A \rightarrow B)}{A \rightarrow B}}{B}}{A \rightarrow B}\,1 \tag{1.53}$$

and then composes

$$\frac{(1.53) \qquad (1.52)}{A} \tag{1.54}$$

This proof of A proceeds according to the pattern Kreisel indicates in his remark (1.4c); the presence of the premiss (1.50) can be seen as a way of expressing within the proof itself that only 'some A' are being considered. Also, because of these premisses, the proof of A does not require us first to prove B, unlike those in sections 1.3.1 and 1.3.3; and unlike those in section 1.3.1, but like those in section 1.3.3, it does require that we have identified the proposition B (or at least the open $B(x)$) when setting out to make inferences from the hypothesis A.

Finally,

$$\frac{(1.54) \qquad (1.53)}{B} \tag{1.55}$$

In the paradoxes at hand, $B = \perp$; in Russell's, moreover $A = a \in a$ for a particular a.[46]

---

[46] That it is possible to have a derivation of 'Russell's antinomy without negation, with exclusive use of the positive propositional calculus' was observed in those terms (in German) by Gödel in *Arbeitsheft* 7, towards the end of 1940: 'Russell Antinomie ohne Negation mit alleiniger Verwendung d. pos. Aussagenkalküls' (Gödel Papers, 5c/19, item 030025, backward direction, p. 7). (A note on p. 12 is marked January 1, 1941.) He labels contraction, in his version $(p \supset (p \supset q)) \supset (p \supset q)$, with 'Entscheidende Formel'. He could have gone on to generalise to the paradox found and published by Curry (1942), but apparently did not.

In the reconstructions of the proof paradoxes below, the use of logic is limited to positive implicational logic (as above), applied to (informally) decidable propositions. While that use of logic is *correct* on the Proof Explanation, it can be treated truth-functionally; from an intuitionistic perspective, one would say that the latter treatment is just another, simpler application of mathematics to the language of mathematics.

There are of course many A and B for which (1.50) holds unproblematically; as Van Benthem (1978, p. 50) reminds us, in his discussion of Löb's Paradox,

$$(A \leftrightarrow (A \to B)) \leftrightarrow A \wedge B \tag{1.56}$$

is a tautology.[47] The present discussion, however, concerns A and B for which the existence of a proof prior to that of (1.50) is not assumed, and where (1.50) is all we know about A – which it is therefore natural to prove as above. (We will see that this is somewhat different in the case of Goodman's Paradox, where we have an A that is naturally proved via a proof of $A \to B$, but where a proof that furthermore includes a proof of $(A \to B) \to A$, while possible, passes via a proof of A, and is therefore not natural.)

There are several ways in which the paradoxes may be avoided by rejecting something in the above derivation. For example, insisting that hypotheses can be used only once, as in linear logic, would make contraction (1.53) impossible; the use of subproofs can be restricted in a way that rules out (1.54) (Fitch 1952, p. 109, and Rogerson 2007 for further discussion); or one may, more vaguely, suggest that we have 'a wrong idea of [the] logical force' of propositions of the form $A \to (A \to B)$ (Geach, 1955, p. 72).[48] Any such choice leads to narrower conceptions of constructive proof than the intuitionistic one which Kreisel was interested in, as were the other originators of the paradoxes discussed below. For them the cause of the problem must lie in the prior way A was concocted.

In contrast to proof paradoxes, provability paradoxes turn not on explicitly formulated properties of a particular proof, but on the existence of any proof whatsoever of a certain self-referential sentence or proposition about provability. Thus they would not serve to illustrate Kreisel's remark (1.4a); but, more importantly for the present discussion, they do not illustrate his remark (1.4c) either.[49] This will be illustrated by Myhill's Paradox (Myhill, 1960, pp. 469–470).[50] Montague's Paradox was originally

[47] Note that a proof from right to left involves an inference from B to $A \to B$ that, on their respective understandings of the Proof Explanation, Heyting would accept, but, arguably, Brouwer would not (in general); see p. 7 above.

[48] Kreisel knew Geach well at the time (letter from Kreisel to Derus, September 3, 2004, in Derus (2020, p. 127)), and aptly connected Geach' paradox (a rediscovery of Curry's Paradox) to Löb's paper when writing about the latter for *Mathematical Reviews* (Kreisel, n.d.-c); my attention to the latter fact was drawn by van Benthem (1978, p. 55).

[49] Independently of Kreisel's remarks, the main question about provability paradoxes is, in a Brouwerian setting at least, whether the sentences or propositions that figure in such paradoxes have any mathematical significance at all (Petrakis (n.d.); Dean (2014, p. 178)).

[50] Gödel's *Arbeitsheft* 16 (Gödel Papers, 5c/28, item 030034) gives on p. 89 what seems to be the origin of Myhill's Paradox; but I do not know the date. (Earlier in the notebook, there are notes on

formulated for necessity instead of provability, but is otherwise the same.[51] Assume that we have a formal system containing primitive recursive arithmetic, an informal provability predicate B on sentences, the reflection axiom schema

$$\vdash \overline{B}(\overline{\ulcorner p \urcorner}) \to p\,, \tag{1.58}$$

and the inference rule

$$\frac{p}{\overline{B}(\overline{\ulcorner p \urcorner})}\;\mathrm{R}\;. \tag{1.59}$$

By the Diagonal Lemma (p. 20 above),[52] there is a sentence p_0 such that

$$\vdash p_0 \leftrightarrow (\overline{B}(\overline{\ulcorner p_0 \urcorner}) \to \bot)\,. \tag{1.60}$$

Myhill derives a contradiction from (1.58)–(1.60) and classical reasoning, but positive implicational logic suffices. Decompose (1.60) into two implications. With contraction provided by (1.53), we then first derive

$$\dfrac{\dfrac{\overline{B}(\overline{\ulcorner p_0 \urcorner}) \to p_0 \qquad p_0 \to (\overline{B}(\overline{\ulcorner p_0 \urcorner}) \to \bot)}{\overline{B}(\overline{\ulcorner p_0 \urcorner}) \to (\overline{B}(\overline{\ulcorner p_0 \urcorner}) \to \bot)}\;\text{transitivity}}{\overline{B}(\overline{\ulcorner p_0 \urcorner}) \to \bot}\;\text{contraction}\;, \tag{1.61}$$

use that in

---

his own Russell paper and on the general theory of relativity.) We here use Gödel's notation. The two axioms and the inference rule that are used for B are as in Gödel (1933a), but compared with that note, here the notation for implication is as the intuitionistic one, and negation as the classical. Instead of a P such that $P \equiv {\sim}B(P)$, which would be the analogue to (1.60) in the text, the P used is such that ${\sim}P \equiv B(P)$. Constructively that does not amount to the same; but, as seen in the justification of step 6 below, Gödel reasons classically.

$${\sim}P \equiv B(P)$$

$$\begin{array}{lll} [0.] & B(P) \supset BB(P) & \\ 1. & {\sim}P \supset B({\sim}P) & \text{aus dem Vorhergehenden wegen } BP = {\sim}P \\ 2. & B(P) \supset {\sim}B({\sim}P) & \text{wegen Ax}\langle\text{ioms}\rangle\text{ für Widerspruchsfreiheit} \\ 3. & B(P) \supset {\sim}{\sim}P & \text{wegen 1 2 und Aussagenkalkül} \\ 4. & B(P) \supset {\sim}P & \text{weil } BP = {\sim}P \text{ und Aussagenkalkül} \\ [5.] & {\sim}B(P) & \text{wegen 3 4 und Aussagenkalkül} \\ [6.] & P & \text{wegen } {\sim}BP = P \\ [7.] & B(P) & \text{wegen Ax}\langle\text{iomensystems}\rangle \end{array} \tag{1.57}$$

In step 2 is meant the axiom $B(p) \supset p$, and in step 7 the inference rule 'From p, conclude to $B(p)$'. In step 3, Gödel first had written P instead of ${\sim}{\sim}P$ and then crossed it out. (Transcription from the Gabelsberger MvA.)

[51] For comparison of the paradoxes of Goodman and Montague, see Weinstein (1983, pp. 264-265), Dean (2014, pp. 164-165, p. 188n17), and Dean and Kurokawa (2016, pp. 40-44).

[52] The Diagonal Lemma can also be shown using the Fixed Point Theorem (see the next section), so in that sense Myhill's Paradox, too, can be seen as an application of the latter.

$$\frac{(1.61) \qquad (\overline{B}(\overline{\ulcorner p_0 \urcorner}) \to \bot) \to p_0}{\dfrac{p_0}{\overline{B}(\overline{\ulcorner p_0 \urcorner})}\,\mathrm{R}}\ , \tag{1.62}$$

and again in

$$\frac{(1.62) \qquad (1.61)}{\bot}\ . \tag{1.63}$$

Although the proof of $\overline{B}(\overline{\ulcorner p_0 \urcorner})$ in (1.62) proceeds via a (0-step) proof of $\overline{B}(\overline{\ulcorner p_0 \urcorner}) \to p_0$, as the latter is a premiss in (1.61), it does not exemplify Kreisel's remark (1.4c), as it contains no subproof of $(\overline{B}(\overline{\ulcorner p_0 \urcorner}) \to p_0) \to \overline{B}(\overline{\ulcorner p_0 \urcorner})$. Nor can it be restructured to that effect, as the rule R cannot be used if there is an open assumption.[53]

1.4.2 Lawvere's Fixed Point Theorem

In the reconstructions of the proof paradoxes below, a central role is played by the following result.

Theorem 5 (Fixed Point Theorem) *(Lawvere, 1969) Let A and B be any objects in a category with a terminal object* 1 *and finite products. Suppose that there exists a morphism $g\colon A \times A \to B$ such that for every $f\colon A \to B$ there exists an $a\colon 1 \to A$ that represents it via g, in the sense that for all $x\colon 1 \to A$, $<a, x>g = xf$. Then for all $h\colon B \to B$ there exists a $b\colon 1 \to B$ such that $b = bh$.*

Proof Let a g and h as described be given. The diagonal morphism $\Delta_A\colon A \to A \times A$ sends $x\colon 1 \to A$ to $<x, x>$. Define the morphism $k\colon A \to B$ as the composition $\Delta_A g h$:

$$\begin{tikzcd} A \times A \arrow[r, "g"] & B \arrow[d, "h"] \\ A \arrow[u, "\Delta_A"] \arrow[r, "k"] & B \end{tikzcd} \tag{1.64}$$

By hypothesis, k is represented by some $a\colon 1 \to A$. Now consider $ak\colon 1 \to B$. By representation of k, $ak = <a, a>g$; by definition of k, $ak = <a, a>gh$. Hence ak is a b as sought. □

[53] A natural and to my mind correct reaction is to say that this reflects a forgetfulness of this formalism with respect to the thought it formalises: we justify the inference rule R (1.59) by pointing to the derivation leading up to its premise, but that derivation is not represented in its conclusion. Artemov (2001) presents a Logic of Proofs (LP) that corrects this, Fitting (2008) extends that into Quantified Logic of Proofs (QLP), and Dean (2014) uses the latter to analyse the Myhill-Montague Paradox. Thereby the provability paradox (which does not mention any particular proof) becomes a proof paradox (which does). Dean and Kurokawa (2016, section 5.5) discuss their reconstruction of Goodman's Paradox in terms of QLP. Note that the Logic of Proofs is not an *analysis* of the Proof Explanation (Artemov, 2001, pp. 2–3).

The following points will be useful for the purposes of the present paper:

1. If the hypothesis of the theorem is satisfied constructively, the conclusion holds constructively.
2. The construction in the proof also goes through if the definition of the $f: A \longrightarrow B$ depends on parameters, so that for example $f_v: A \longrightarrow B$ leads to a fixed point b_v. (This plays a role in the reconstruction of Goodman's Paradox in section 1.4.4.4.)
3. If A and B are objects in a category where the objects are collections, the existence of a morphism $a: 1 \longrightarrow A$ corresponds to truth of the proposition $a \in A$, and the composition af to the application $f(a)$. This is the case everywhere in the present paper.[54]
4. The definition of $f: A \longrightarrow B$ is presupposed in that of a representation a of it, whence the latter definition is impredicative, as a lies in the range of arguments of f. A predicative characterisation of provably the same object may or may not be simultaneously available. If not, then this definition is 'critically impredicative'. The use of the term 'critical' here comes from Bernays (1962): it is such cases that raise the question whether the definition is constructively acceptable. (See section 1.4.5 for a quotation and further discussion.)
5. The condition on g has the form $\forall f \exists a R(f, a)$. This entails that there be a choice operation F, i.e., $\exists F \forall f R(f, F(f))$, and we can set $a = F(f)$.[55] F is typed as $(A \longrightarrow B) \longrightarrow A$; it is a selection functional. The verification that the element $a = F(f)$ is of type A corresponds to a proof of the proposition A from proofs of the propositions $A \longrightarrow B$ and $(A \longrightarrow B) \longrightarrow A$. We saw a similar situation when discussing proof skeleton (1.15). It also presents a contrast: there, an element b of B was used towards constructing a selection functional; here, it is the reverse.
6. The condition on g is, in asking for a representation of every $f: A \longrightarrow B$, stronger than required for the proof, which uses only representability of morphisms $A \longrightarrow B$ composed like k, with varying h. This was observed in Yanofsky (2003, p. 378), and is exploited in the proof of the Diagonal Lemma for Terms just below.[56]

[54] In a specifically intuitionistic context, such objects would be either species or denumerably unfinished sets. For the respective definitions, see (1.125) and the comment on (1.124) below.

[55] The intuitionistic justification of choice principles depends on their being construed intensionally; results to the effect that choice implies the Principle of the Excluded Middle (Diaconescu, 1975; Goodman & Myhill, 1978) require an extensional construal (Martin-Löf, 2006).

[56] As I was finalising the present paper, my attention was drawn (by Andrei Rodin) to the preprint Roberts (2021). Following up on Yanofsky's observation, there the general question is answered what the weakest hypothesis is that allows to prove the conclusion of the Fixed Point Theorem. Roberts shows that (*i*) a more general notion of product than the standard one suffices (which enables him to state, in Example 1, a fixed point result of the same kind that would have been missed under the original hypothesis), and (*ii*) quantification over all or certain morphisms $f: A \longrightarrow B$ can be done away with. However, the conclusion is still dependent on the presence of an element of A that represents one particular morphism on all of A. (See the hypothesis of his Theorem 14, and the gloss on it: 'if there is some $\mathsf{a}_0 \in \mathsf{A}$ such that $\mathsf{F}(\mathsf{a}_0, -) = \sigma \circ \mathsf{F} \circ \delta_{\mathsf{A}}$', where the placeholder is one for elements of A.) That is crucial to my discussion of the Proof Paradoxes in section 1.4.5, as this means that the impredicativity noted in point 4, on which my diagnosis will depend, remains.

7. The proof of the theorem in effect constructs a choice operation H and puts $b = H(h)$. Thus, $\forall h \exists b(b = bh)$ is proved via $\forall h \exists b(b = bh) \leftrightarrow \exists H \forall h(H(h) = H(h)h)$ and a proof of the right hand side. This yields an overall proof with skeleton (1.16).

The existence of a representation of each $f : A \to B$ via an appropriate g, required by the hypothesis of the theorem, will have to be justified by an appeal to some general principle. In Russell's Paradox, that role is played by unrestricted comprehension. In the other paradoxes analysed below these are principles of the form:

If we have a performable operation (act) for assigning a construction (object) to another construction (object), then to this operation corresponds a function(al), which is itself a construction (object).[57] (1.65)

The recognised operation serves to define the function(al). Thus, in a category where the object A is the class of all constructions, and morphisms are performable operations, the morphisms $A \to A$ are themselves very much like elements of A, and can be identified as such. The resonance of (1.65) with the introduction to Gödel's *Dialectica* paper, with its emphasis on 'Denkgebilde' (1958, p. 280), is intended: the concern here is with constructivity as a mode of mental operation, reflection on which leads to the introduction of constructions as objects. Since the perspective from which I look at comprehension and principles of the form (1.65) is, for the present purpose, determined by their role in applications of the Fixed Point Theorem, I will use the umbrella term 'representation principles'.

A common alternative to the Fixed Point Theorem in presentations of paradoxes is the fixed point combinator $Y = \lambda y.(\lambda x.y(xx))(\lambda x.y(xx))$, which has the property $Yz \equiv z(Yz)$ for any term z. An apposite example in the present context is the reconstruction of Goodman's Paradox in Dean and Kurokawa (2016, p. 43). But by itself the existence of such a fixed point term Yz is a syntactical phenomenon in a formal theory; for an analysis of paradoxes in contentual mathematics, which is where they are problematic and where they must be resolved,[58] an approach using the Fixed Point Theorem is more direct, because it shows how fixed points are brought about in contentual terms.[59]

I will end this section with an aside that illustrates of point 6; it is not part of the discussion of the paradoxes. It is a proof of the Diagonal Lemma for Formulas that

[57] Here I am referring to the threefold distinction thematised in Sundholm (1983, p. 164) between construction as (1) a process as it unfolds in time; (2) an object obtained as the result of such a process; (3) a construction-process as object (the objectification of a process of construction). For each one can furthermore distinguish between types and tokens (van Atten, 2018, pp. 1596-1597).

[58] To my mind, the suggestion favourably brought up in Dean (2020, p. 585) that paradoxes might be *resolvable* by mapping them to undecidability results in certain formal theories involves a shift of topic.

[59] Incidentally, the combinator Y can be obtained by putting the Theorem to metamathematical use: Take for A and for B the set Λ of λ-terms; define $g = \lambda x.D_2x(D_1x)$, where the D_i are projection operators; let the representation principle be: morphisms $\Lambda \to \Lambda$ are themselves given by λ-terms; and let h_y be the family of morphisms $\lambda x.yx$. For each h_y, this yields a fixed point $(\lambda x.y(xx))(\lambda x.y(xx))$, and now abstraction on y gives Y. (This derivation has been adapted from Li (2021), were also some related fixed point combinators are derived.)

uses the proof of the Fixed Point Theorem while requiring representability only of k; compared to the version proved in section 1.3.3, this one is for a theory with a richer language. The proof takes the form of a corollary of a Diagonal Lemma for Terms.

Theorem 6 (Diagonal Lemma for Terms) *(Jeroslow, 1973).*[60] *Let S be a system that contains primitive recursive arithmetic and has symbols for all primitive recursive functions. Then for every formula $\varphi(x)$ there is a closed term t such that $S \vdash t = \ulcorner \varphi(t) \urcorner$.*

Proof (adapted so as to use the Fixed Point Theorem). In the category of sets, let A be the set of Gödel numbers of the symbols in S for all primitive recursive functions $\mathbb{N} \to \mathbb{N}$. Let B be the set of Gödel numbers of closed terms in S. Define

$$\begin{aligned} g\colon A \times A &\to B \\ <\ulcorner x \urcorner, \ulcorner y \urcorner> &\mapsto \ulcorner x(\ulcorner y \urcorner) \urcorner \end{aligned} \tag{1.66}$$

and

$$\begin{aligned} h\colon B &\to B \\ \ulcorner t \urcorner &\mapsto \ulcorner \varphi(t) \urcorner \end{aligned} \tag{1.67}$$

Following the proof of the Fixed Point Theorem, $k\colon A \to B$ is defined as $\Delta_A g h$, which has the action $\ulcorner x \urcorner \mapsto \ulcorner \varphi(x(\overline{\ulcorner x \urcorner})) \urcorner$. This action can also be effected by a certain primitive recursive function a, in the language of S symbolised by $\overline{a}$, and with $\ulcorner \overline{a} \urcorner \in A$. The latter therefore serves to represent the morphism k via g.[61] We now have the fixed point $b = \ulcorner \overline{a} \urcorner k = bh = \ulcorner \varphi(\overline{a}(\overline{\ulcorner \overline{a} \urcorner})) \urcorner$, whence $a(\ulcorner \overline{a} \urcorner) = \ulcorner \varphi(\overline{a}(\overline{\ulcorner \overline{a} \urcorner})) \urcorner$, and $\vdash \overline{a}(\overline{\ulcorner \overline{a} \urcorner}) = \overline{\ulcorner \varphi(\overline{a}(\overline{\ulcorner \overline{a} \urcorner})) \urcorner}$. □

In this proof and in that of similar metamathematical theorems, the set A can be defined predicatively, so that the definition of a representation of k, while impredicative, is not critically impredicative.

Corollary 1 (Diagonal Lemma for Formulas) *For every formula $\varphi(x)$, there is a formula ψ such that $\vdash \psi \leftrightarrow \varphi(\overline{\ulcorner \psi \urcorner})$.*

Proof Apply the Diagonal Lemma for Terms to $\varphi(x)$, and set $\psi = \varphi(\overline{a}(\overline{\ulcorner \overline{a} \urcorner}))$; then

$$\begin{aligned} &\vdash \psi \leftrightarrow \varphi(\overline{a}(\overline{\ulcorner \overline{a} \urcorner})) && \text{by def. } \psi \\ &\vdash \psi \leftrightarrow \varphi(\overline{\ulcorner \varphi(\overline{a}(\overline{\ulcorner \overline{a} \urcorner})) \urcorner}) && \text{by the Diagonal Lemma for Terms} \\ &\vdash \psi \leftrightarrow \varphi(\overline{\ulcorner \psi \urcorner}) && \text{by def. } \psi \end{aligned}$$

□

[60] Jeroslow (1973, p. 360) mentions that it was his referee who had isolated this lemma from one of the proofs in Jeroslow's manuscript. Santos (2020, pp. 26, 38) shows that the Diagonal Lemma for Formulas does not, in turn, entail that for Terms.

[61] By point 6 on p. 28, it suffices here to have a representation principle for morphisms $A \to B$ that are composed like k and that are effectible by a primitive recursive function.

1.4.3 Russell's Paradox

This paradox (Russell, 1903) was one of the example reconstructions in Lawvere's paper on his Fixed Point Theorem (1969, p. 137). It is included here (*i*) to enable a direct comparison of the proof paradoxes below with this well known one, and (*ii*) as part of the background to the paradox devised by Kreisel, who actually made that comparison (as seen in (the discussion of) his (1.79)–(1.80b) below.)

For a set-theoretic version, let A be the universe of all sets, and B the set of truth-values $\{\top, \bot\}$. Define

$$\begin{aligned} g\colon A \times A &\to B \\ <y, x> &\mapsto \begin{cases} \top & \text{if } x \in y \\ \bot & \text{if } x \in y \to \bot \end{cases} \end{aligned} \tag{1.68}$$

According to the unrestricted comprehension principle, for every formula with one free variable $\varphi(x)$, there exists the set $\{x \in A \mid \varphi(x)\}$. Hence, for arbitrary $f\colon A \to B$, there exists the set $\{x \in A \mid f(x) = \top\}$. That set represents f via g.[62] If f is given to us as the characteristic function of a predicate P, then that set is definitionally equal to $\{x \in A \mid P(x)\}$.

Now apply the Fixed Point Theorem, taking for h the negation function

$$\begin{aligned} h\colon B &\to B \\ \top &\mapsto \bot \\ \bot &\mapsto \top \end{aligned} \tag{1.69}$$

Then k is the characteristic function of the predicate $P(x) = x \in x \to \bot$, and is represented by the set a of all sets that do not contain themselves. Evidently, the impredicativity in the definition of a is critical. The conclusion $b = bh$ here means, in propositional terms, that $a \in a \leftrightarrow (a \in a \to \bot)$. Now one one derives $\bot$ as in (1.50)–(1.55).

For a property-theoretic version, let A be the universe of all objects, including properties, and replace $\in$ by the exemplification relation ϵ. The unlimited abstraction axiom states that for every unary predicate P the property $\lambda x.P(x)$ exists. With the Fixed Point Theorem, we find the property $a = \lambda x.(x \mathrel{\epsilon} x \to \bot)$ of being not self-exemplifying, for which $a \mathrel{\epsilon} a \leftrightarrow (a \mathrel{\epsilon} a \to \bot)$.

Russell (1906, pp. 35–36) offered a generalisation of the set-theoretical version; it is this generalisation that Kreisel appeals to in (1.80a) below. In the present terms: Let A be the universe of all sets, u, x, and y variables ranging over A, and $j\colon A \to A$ a morphism such that

$$\forall u(\forall x(x \in u \to P(x)) \to \exists y(y = j(u) \wedge P(y) \wedge (y \in u \to \bot)))\,. \tag{1.70}$$

---

[62] The circumstance that here the set is seen as a representation of its characteristic function, instead of the more usual converse view, is of course a consequence of the category-theoretical view of an element of a set as a morphism from a terminal object.

Russell calls a set u for which the condition holds, as well as the process of applying j to it, 'self-reproductive', in that the result of this application is again a set for which the condition holds (Russell, 1906, p. 36). Define

$$\begin{aligned} g\colon A \times A &\to B \\ <y,x> &\mapsto \begin{cases} \top & \text{if } j(x) \in y \\ \bot & \text{if } j(x) \in y \to \bot \end{cases} \end{aligned} \tag{1.71}$$

By unrestricted comprehension, for every $f\colon A \to B$, there exists the set $\{z \in A \mid \exists x(z = j(x) \wedge f(x) = \top)\}$. That set represents f via g. Taking for h the negation mapping, we conclude to the existence of a set a such that $j(a) \in a \leftrightarrow (j(a) \in a \to \bot)$. The earlier version is the special case where j is the identity map and $P(x)$ is $x \in x \to \bot$. For let u be any set for which $\forall x(x \in u \to (x \in x \to \bot))$. Instantiating x with that u and contracting as in (1.53) yields $u \in u \to \bot$, from which each of the three conjuncts within the scope of the existential quantifier in (1.70) follows immediately.

1.4.4 Proof paradoxes

1.4.4.1 Kreisel's Paradox

This paradox appears as part of Kreisel's second presentation of his Theory of Constructions in his 'Mathematical logic' (1965), in a volume edited by Saaty. The aim of that theory, he explained in the first presentation, was to give 'a formal semantic foundation for intuitionistic formal systems in terms of the abstract theory of constructions' (Kreisel, 1962a, p. 198), beginning with the logic. (The Theory of the Creating Subject, which figures in section 1.4.4.3 below, was meant to be developed into an extension.) The value of the exercise was seen not to lie in conveying the meaning of the intuitionistic constants, but in technical applications such as independence proofs. As Kreisel related to Heyting in a letter of October 5, 1961, Gödel saw a further interest:

(1.72) Gödel regards this whole work as specially interesting from the point of view of the paradoxes. For, on the one hand we use constructions without type distinction, on the other, we avoid paradoxes by not allowing propositions as mathematical objects. (The rules of proof used in the antinomies are intuitionistic: the question is why $X \in X$ cannot be expressed. Type distinctions are certainly not <u>always</u> observed, e.g. not in your explanation of the logical constants, in particular of implication.) (Heyting Papers, Bkre 611005)

A version of that last sentence is included in the published paper (Kreisel, 1962a, p. 202), but not of the rest of this remark.[63]

Kreisel's Paradox, as I call it here, is distinct from the paradox that has become known as the 'Kreisel-Goodman Paradox', which is only found in Goodman's writings (on his modification of Kreisel's theory),[64] and will be discussed in section 1.4.4.4. The present paradox arises upon Kreisel's introduction of two of the main ingredients of his Theory of Constructions, unless further precaution is taken, which he of course goes on to do. The one is 'notions, that is, understood, decidable properties of mathematical objects', where is not required that such a decision can be mechanised (Kreisel 1965, pp. 2.13, 2.14). The other is a convention that, in appropriate contexts, ensures totality of functions, in the interest of having a theory with decidable equality. It consists in a modification of the ordinary meaning of application:

> 2.15. *The Meaning of the Primitive Concepts.* It is not to be expected that we have a clear idea of the *extension* of the concept of mathematical object; [. . .] But the application operation requires a careful interpretation. (1.73)
>
> 2.151. *Total functions.* [. . .] [If] for the objects a, b as *given* or *conceived*, no sense is assigned to $a(b)$, then $a(b)$ is put $= a$, say. (Cf. in type theory: if no sense is assigned to $a \in b$, we regard $a \in b$ as false.) Obviously, for any proposed axiomatic scheme one has to verify its validity for this convention. This is illustrated by considering
>
> 2.152. *The λ-Calculus.* The naive proposal (parallel to the principle: every property defines a collection) is this. For every term $t[x]$, built up by means of the application operation from constants and containing the variable x, there is a function $\lambda z t[z]$ for which we have a proof $a_t : x \cdot (\lambda z t[z])(x) = t[x]$. This is excluded by the *paradoxes*. There is a notion η, $\eta(x) = 0$ if $x \neq 0$, $\eta(x) = 1$ if $x = 0$; it is a notion since any clearly conceived object either is conceived as 0 or not. Consider the term $\eta(x(x))$; though, by the convention of 2.151, it is well defined for each x, is there a clearly conceived object c (c: for Church) with $c(x) = \eta(x(x))$? No, since $c(c)$ and $\eta(c(c))$ are different. In short, the existential assumptions implicit in unrestricted λ-term formation and conversion are not correct. The 'rule' c: for each x, take the value $\eta(x(x))$ overlooks the tacit convention that, for $x = c$, the value is also $c(c)$. (Kreisel, 1965, pp. 124–125, original italics)

Upon reading 2.151, to some it will occur to ask: What if '$a(b)$' takes on a sense only at a certain point in time? Or what if the sense changes? But one just looks at the

[63] I have not found a reply by Heyting in the Heyting Papers or the Kreisel Papers. The Heyting-Kreisel correspondence began in 1952 with a letter from Kreisel, and seems to have ended (with a bang) in 1970; Heyting passed away in 1980. Given that Kreisel attended a lecture by Brouwer as early as 1946 (Kreisel, 1987b, pp. 146), it is regrettable that, for all we know, he did not also begin a correspondence with him (Brouwer died in 1966). Would the reason really just be that Kreisel did not like Brouwer and his style (Kreisel, 1987b, pp. 146-147)? Two ideas of Brouwer's that are not prominent in Heyting's thinking but that Kreisel had a lively interest in are that of proofs as infinite objects (see footnote 36 above) and Creating Subject arguments.

[64] On this last point, see Dean and Kurokawa (2016, p. 40).

situation at the moment that the application is attempted. Time is made an explicit parameter in one version of Gödel's Paradox and in Troelstra's Paradox, discussed in sections 1.4.4.2 and 1.4.4.3 below.

In his review of Kreisel's paper for the *Journal of Symbolic Logic*, Vesley wondered 'whether the system of 2.152 is intended to be a description of calculus of λ-conversion of Church XVII 76 (if so it is a misunderstanding and, in any case, the attribution of the object c to Church is in error)'.[65] The perceived misunderstanding will have turned on the contrast between Church's calculi – 'so formulated that it is possible to abstract from the intended meaning' (Church, 1941, p. 1) – and Kreisel's notions, which are properties that are *understood*. (Church 1941 is the one entry by Church in Kreisel's list of references.) For the paradox as such, this does not matter, as we can view the role of the λ-calculus here as a superstructure; and this agrees better with Church's earlier system of 1932–1933, with its underlying 'intuitive logic' and postulates (Church, 1932, section 4). Whether the attribution of the object c to Church is justified will be addressed in a moment.

To see Kreisel's Paradox as an application of the Fixed Point Theorem, note that his proposal at the beginning of 2.152 can take the role of a representation principle. Kreisel does not introduce an explicit application function, but it seems faithful to 2.15 and 2.151 to define

$$\begin{aligned} g\colon\ & A \times A \to A \\ & < j, x > \mapsto \begin{cases} j(x) & \text{if for } j \text{ and } x \text{ as given or conceived,} \\ & \text{a sense is assigned to } j(x) \\ j & \text{otherwise} \end{cases} \end{aligned} \tag{1.74}$$

where A is the universe of constructions. The case distinction is decidable (as above: not necessarily by a mechanism, but by anyone who understands the meaning of the terms involved), hence g is everywhere defined: that is, even though there is no construction method for its domain, we see that, whenever we have made two constructions j and x, we are able to evaluate g.

For h we take Kreisel's η:

$$\begin{aligned} h\colon\ & A \to A \\ & x \mapsto \begin{cases} 0 & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases} \end{aligned} \tag{1.75}$$

with the understanding that, since η is a notion, h is constructive. Now the Fixed Point Theorem applied to g and h yields a morphism $k\colon A \to A$ whose composition is that of Kreisel's $\eta(x(x))$, i.e., of $h(g(<x, x>)$. By the chosen representation principle, this morphism is represented by a construction of which we have a proof that it is an everywhere defined function, and which, by reflection, therefore is one. This function is Kreisel's c. As c lies in its own domain, self-application makes sense, and we have $g(c, c) = h(g(c, c))$. But that is impossible; in terms of propositional

[65] The Vesley file in the Kreisel Papers (21/6) contains no material related to this review.

logic,

$$c(c) = 0 \leftrightarrow (c(c) = 0 \to \bot)\,. \tag{1.76}$$

Therefore, the function c does not exist, and neither does the proof that it is total: the 'existential assumption' invoked in 2.152 had given *both*. Note that the definition of c is impredicative, and critically so, there being no construction method for the elements of the universe of constructions.

In the version of his Theory of Constructions that Kreisel goes on to develop, he keeps the convention 2.151, but introduces in 2.22 a restriction on types in λ-abstraction that renders the representation principle invalid.[66]

The formal resemblance of Kreisel's Paradox to Russell's is of course clearest from their reconstructions in terms of the Fixed Point Theorem, and hinted at in less general terms by Kreisel's indication, at the beginning of 2.152, of the parallelism between the function-forming principle formulated there and the principle that every property defines a collection: it was in terms of that latter principle that he had presented Russell's Paradox earlier on in the same paper (see (1.80a) below).[67] But of particular interest is to juxtapose 2.151 and 2.152 to the presentation of Russell's Paradox in Church's unpublished Princeton lecture notes 'Mathematical logic' of 1935–1936, which runs as follows. There exists a propositional function $\lambda x.{\sim}x(x)$. Consider the self-application $P = (\lambda x.{\sim}x(x))(\lambda x.{\sim}x(x))$. Then P converts to ${\sim}P$ and vice versa, from which it would seem to follow that P is a proposition that is both true and false.

> The conclusion to be drawn is that the range of the independent variable of some ppfns [i.e., propositional functions] is not and cannot be the universe. For, in the case of the ppfn $\lambda x.{\sim}x(x)$, the Russell paradox would result, not only from the assumption that the range of the independent variable of this ppfn was the universe, (1.77a)

[66] I take it that when Kreisel in his earlier presentation of the Theory of Constructions voiced doubts about the consistency of one of its variants (Kreisel, 1962a, pp. 200, 203), what he saw coming ahead was a paradox like the one he came to formulate in 2.152. Related to this: In correspondence with Goodman (Kreisel Papers, 50/3), Kreisel notes that '${\sim}\pi(p_0, x, \top)$ has the properties of Church's η, $\underline{\underline{\eta(x) \neq x}}$'. Here p_0 is a construction that proves the extensional equality of $\top$ and $\top$ (as defined in a manuscript of Goodman that Kreisel is commenting on), and $\pi(a, b, c)$ is the two-valued term defined in Kreisel (1962a, p. 203), interpreted as 'the construction a is a proof that b and c are extensionally equal'. It is not clear to me whether Kreisel means to ascribe η to Church or only to associate it with him, but am inclined to believe the latter, as I have not found it in Church. In that same folder lies a note (in the handwriting Kreisel used when writing to himself) relating a paradox of the self-application of $\lambda x.\pi(p, {\sim}xx, \top)$ to c and η as in 2.152 (but without that reference). The note is not dated, but the page also contains typewritten corrections to what seems to be a version of the Saaty paper.

[67] As far as the presentation of these paradoxes is concerned, one notes that 2.152 and the unnumbered presentation of Russell's Paradox on pp. 100–101 contain no explicit reference to one another.

(1.77b) but equally from the assumption that it was possible to define a ppfn f such that $f(a) = \{\lambda x.{\sim}x(x)\}(a)$ whenever the expression $\{\lambda x.{\sim}x(x)\}(a)$ had a meaning and $f(a) = F$ in all other cases. (Church, 1935–1936, p. 17)[68]

The possibility for generating Russell's Paradox in (1.77b) is not mentioned in Church's other presentations (1932, p. 347; 1933, pp. 860–861; 1941, pp. 70–71). His solution is to say that P is not meaningful because it has no normal form (Church 1935–1936, p. 17; 1941, pp. 15, 70–71), and the same could be said about $f(f)$. But one now notes that Kreisel, in effect and perhaps in intention, regains the paradox in (1.77b) by generalising propositional negation to a notion that is meaningfully applied to any object and that, to put it anachronistically, likewise has no fixed points.[69] The role of Kreisel's function c is analogous to that of Church's f, and Kreisel's convention and Church's assumption make these functions total in the same manner. In this sense, while attributing c to Church would be mistaken, there would be room for associating it with him. (As for η, see footnote 66.) Coming to circumstantial evidence, we note that Kreisel wrote the Saaty paper while in Princeton (Kreisel, 1965, p. 191), so he may have discussed the matter with Church (although no conversation with him is acknowledged), or at least have read Church's lecture notes there.[70]

Kreisel gives a diagnosis of his Paradox in a letter to Gödel of April 1, 1968.[71]

(1.79) Vielen Dank für die angenehmen und nützlichen Gespräche. Es ist nur schade, daß ich bei unserer Diskussion der Church'schen Paradoxie eine, m.E. wesentliche, Unterscheidung nicht genügend betont habe, nämlich zwischen *Verfahren* und *Funktion* (oder Konstruktion).

Eine Funktion f ist ein Tupel (Verfahren V_f, Menge D_f; Einsicht E_f daß V_f auf D_f definiert ist). Noch etwas präziser: Verfahren führen immer (sozusagen hereditär) von Verfahren + Definitionsmenge zu Verfahren + Definitionsmenge einerseits, von Einsichten zu Einsichten andererseits. Man vermischt nicht das Objektive und Subjektive.

---

[68] Kreisel of course also knew this passage in Gödel, which echoes Church's:

(1.78) It should be noted that the theory of types brings in a new idea for the solution of the paradoxes, especially suited to their intensional form. It consists in blaming the paradoxes not on the axiom that every propositional function defines a concept or class, but on the assumption that every concept gives a meaningful proposition, if asserted for any arbitrary object or objects as arguments. The obvious objection that every concept can be extended to all arguments, by defining another one which gives a false proposition whenever the original one was meaningless, can easily be dealt with by pointing out that the concept 'meaningfully applicable' need not itself be always meaningfully applicable. (Gödel, 1944/1951, p. 149)

[69] The presentation of Russell's Paradox in (Curry, 1934, pp. 588–589) uses N instead of $\sim$, and a predicate Pr that is true of propositions; thus even more readily suggesting their reinterpretation or replacement by something more general.

[70] The Church file in the Kreisel Papers (3/6) contain no material related to this.

[71] A date in the middle of the period during which he was working on the Brouwer obituary.

Es gibt natürlich Verfahren, die sich nur auf Verfahren (ohne Erwähnung der Definitionsmenge des Arguments) beziehen, z.B. Konstante.

CHURCH'SCHE PARADOXIE (siehe 2.151, 2.152 auf S.124–125 des Artikels „Math. Logic", in Saaty). Natürlich „beschreibt" $\eta(x(x))$ ein Verfahren, worin die Definitionsmenge D_x *selbst* eingeht (von einem *Verfahren* V_x kann man nicht entscheiden, ob V_x für das Argument V_x definiert ist und S.125, Z.3–5, entsprechen keinem Verfahren).[72] Nennen wir dieses Verfahren V_c, wie im Saaty Band.

Ehe man zur Church'schen Paradoxie kommt, muß man noch ein D_c angeben, von dem man weiß, daß V_c auf D_c definiert ist (also eine Einsicht E_c, die von D_c abhängt).

Alles, was die Church'sche Paradoxie zeigt, ist m.E. dies: Obwohl wir *ein Verfahren* V_c haben, haben wir mehrere Funktionen, abhängig von D_c. Und (V_c, D_c) ist nicht in D_c: V_c ist also auch auf $D_c \cup (V_c, D_c)$ definiert.

Ist das nicht analog zu Saaty, S.100-101 und Fußnote 4 auf S.101?

Beste Grüße, auch an Ihre Frau (der es hoffentlich besser geht)

Ihr sehr ergebener G Kreisel

(Kreisel Papers, 50/1)

'Church's Paradox' is also the name Gödel used for a paradox to be discussed in the next section. Since differences between the paradoxes as presented by Kreisel and Gödel are important here, and these two paradoxes were, in their specifics, not devised by Church, I will not use the name 'Church's Paradox' for either.

As Kreisel views things here, a 'Verfahren' is an operation that can be carried out on one or more objects supposed to have been given, and is expressed in a rule with free variables; unlike a function, it is given without a domain. The paradox is then used to show that while the operation implicit in the notion η is applicable to any construction, there is no function (or construction) to the same effect.

At the end, Kreisel compares this to Russell's Paradox, by referring to the pages where he had given the latter in a form of the self-reproductivity argument (see (1.70) above):

> [T]here is a genuine problem: what properties define collections, particularly if properties themselves are to be regarded as objects. This may be shown by means of the *paradoxes*. If x is the collection of objects satisfying a property, take $\{x\}$ (whose only element is x) to be the property regarded as an object. Let r (for Russell) be *any* collection satisfying $(\forall x)(x \in r \Rightarrow x \notin x)$, and so $r \notin r$. Thus, $\forall x(x \in r \cup \{r\} \Rightarrow x \notin x)$, $r \cup \{r\} \supset r$ but $r \cup \{r\} \neq r$. In other words, r is not the collection of all objects satisfying $x \notin x$. (1.80a)

The 'footnote 4' Kreisel refers to at the end of (1.79) is to this passage and reads:

> One of the set theoretic *definitions* of ordinals takes the empty set $\emptyset$ to be zero: $a \to a \cup \{a\}$ as the successor function (and unions for limits). The (1.80b)

[72] [Note MvA] The lines referred to occur in 2.151 quoted as (1.73) above: '[If] for the objects a, b as *given* or *conceived*, no sense is assigned to $a(b)$, then $a(b)$ is put $= a$, say.'

> argument above is literally the proof that there is no greatest natural number (greatest ordinal). (Kreisel, 1965, pp. 100–101, original emphasis)

Or, as he put it in his earlier presentation of this analogy,

(1.81)
> From this point of view the Russell Paradox does not seem more astonishing than a child's assumption that there is a greatest integer: we have overlooked the fact that not every property has a definite extension. (Kreisel, 1958, p. 157)[73]

The analogy, then, is that just as collections that have the property 'being a collection of elements that do not contain themselves' or 'being a collection of ordinals closed under predecessor' are self-reproductive, so are collections that have the property 'being a domain on which V_c is defined'; hence, just as not every property has a definite extension, not every operation (*Verfahren*) determines a function (construction).

I have not found a letter or note by Gödel in which he explicitly sets out to give an answer to Kreisel's letter. Perhaps there was one and it got lost. Either way, the matter will surely (also) have been discussed on one of their phone calls, or on one of Kreisel's visits to Princeton. In fact, Kreisel has said that 'My main contact with Gödel was in private conversations during the years I spent at the same Institute, not in correspondence'.[74] A plan for one such visit was mentioned by Kreisel in between that letter and the Buffalo conference, held in August 1968.[75] Be that as it may, we will now see that Gödel had a different view, and that Kreisel came to change his mind.

1.4.4.2 Gödel's Paradox (intuitionistic)

There is a paradox associated with Gödel that comes in both a classical and an intuitionistic version; the latter is a proof paradox. They are reproduced here from Gödel's archive in Fig. 1.1 and Fig. 1.2. These notes were written back to back, and kept in an envelope marked 'Antin⟨omien⟩ des Intuit⟨ionismus⟩ und der abs⟨olute⟩ Beweisbarkeit'.[76] (Another intuitionistic paradox in that envelope is reproduced in Fig. 1.3.) Gödel named the classical version 'Church's Paradox', 'because it is most easily set up in Church's system' (Wang, 1996, p. 279, 8.6.24).[77] Wang proposes to call it 'Gödel's Paradox'; for the reason given on p. 37 above, so will I.

I do not know when Gödel first thought of either version, which of the two came first, or how much time there was between them; but I will suggest that at least the

[73] We will come back to this remark on p. 67.

[74] Letter of November 2, 2004 to Kai Käkelä, quoted in Derus (2020, p. 128).

[75] In a letter from Kreisel to Gödel, July 11, 1968 (Kreisel Papers, 50/3).

[76] A pencil note on it states that it was 'filed with Wang corresp⟨ondence⟩'.

[77] Gödel reviewed Church (1932) and Church (1933) for the *Zentralblatt*, but without comments (1932; 1934). The inconsistency referred to at the end of Gödel (1934) will have been the one established by Kleene and Rosser in the Spring of 1934 – Gödel was in Princeton from October 1933 through May 1934 (Kleene 1981, p. 57; Wang 1987, p. 95).

particular way in which Gödel presents the intuitionistic version in Fig. 1.2 is a direct reaction to Kreisel's Paradox. In the previous section, it was remarked that Kreisel may have arrived at his paradox by generalising Church's quoted in (1.77b) above. That route was, of course, at any point open to Gödel as well. (For the principal purpose of illustrating the 'errors' referred to in Kreisel's remark (1.4a), obviously neither the considerations in the present paragraph, nor similar ones below, matter. But they are motivated by a perceived intrinsic interest, and by the idea that they may lead to clues as to what, in the sources, should be read in the light of what.)

Gödel never published his paradox. Kreisel included the classical version, in a slightly different presentation, in the proceedings of the Logic Colloquium '69, calling it 'a standard "functional" paradox', without an attribution to Gödel (Kreisel, 1971b, pp. 190-191). Wang included it in his *Logical Journey*, based on his conversation with Gödel of October 18, 1972 (Wang, 1996, p. 278–279). No mention is made of Kreisel.[78] The intuitionistic version is referred to in Wang's book (p. 279), but not included.

Bew⟨eis⟩ (0.):

Df. von E:

$$E(x) = 0 \text{ wenn } x \neq 0$$
$$E(0) = 1 \qquad \text{dann } \underline{E(x) \neq x}.$$

$a = b$ bedeutet: *a is the object b*

Church Antinomie

Df. 0. Überall definierte Funktion. sind überall definiert
Th. 1. Es gibt überall definierte Funktionen, zum Beispiel $I(x)$, $E(x)$, = (von 2 Variablen).
Df. 2. $\underline{F\,.\,x = F(x)}$ wenn F eine überall definierte Funktion ist, sonst = 0.
alle
Th 3. . ist eine überall definierte Funktion und $\underline{F\,.\,x =_x F(x)}$ für überall definierte Funktion⟨en⟩ F.
Df 4. $\underline{H(x)} = E(x\,.\,x)$. H ist ⟨eine⟩ überall definierte Funktion.
Th. 5. $\underline{H(x)} = H\,.\,x = E(x\,.\,x)$
6. $H\,.\,H = E(H\,.\,H)$
andererseits " $\neq$ "

Fig. 1.1 Gödel's Paradox, classical (Gödel Papers, 12/52, collective item 060772). (Transcription from the Gabelsberger MvA, advised by Robin Rollinger.)

The paradox is, Gödel comments, 'a simpler version of the familiar paradox of the concept of not applying to itself' (Wang, 1996, p. 279, 8.6.24) – the property-

[78] That may of course be an artefact of Wang's representation. But in view of the date, one can't help recalling the fact that the to all appearances last letter from Kreisel to Gödel was dated October 1, 1972. In that letter, Kreisel had expressed some grievances against Gödel, and reflected that 'Nach meiner Erfahrung mit vielen Menschen (nicht aufgrund ungeprüfter "Theorien") scheinen die menschlichen Beziehungen eine gewisse natürliche Lebensdauer zu haben'. As Parsons comments, 'It is very probable that they never saw each other again' (Parsons, 2020, pp. 79, 83).

<u>*Antin⟨omie⟩* im *Intuit⟨ionismus⟩*</u>

Df <u>überall definierte Funktion</u>[x] (unentscheidbar)
+ Beispiele

<u>Df</u> Wenn f ein Paar $< B, g >$, B ein Beweis ist dass das Verfahren g überall zum Resultat führt:

$f\,.\,x =_{Df} g(x)$
sonst $=_{Df}$ 0   *<u>Th</u>* . ist überall definiert.

(0.) Es gibt ein nachweislich überall definiert⟨es⟩ E so, dass $E(x) \neq x$ (siehe Rückseite).
(1.) Es gibt ein H so, dass:

$$H\,.\,x = E(x\,.\,x).\text{ Dann ist:}$$
$$H\,.\,H = E(H\,.\,H)$$
$$\text{andererseits } H\,.\,H \neq E(H\,.\,H)$$

<u>*Bew*⟨eis⟩ (1.)</u>:
Es gibt ein überall definiertes[a] Verfahren G so, dass $G(x) = E(x\,.\,x)$. Also: es gibt einen Beweis B der zeigt, dass G ein überall definiertes Verfahren ist. $H =_{Df} < B, G >$. Dann: $H\,.\,x = G(x) = E(x\,.\,x)$.

[x]Funktion = Operation = Verfahren (Regel der Verwertbarkeit = Erstellung einer Reihe von Gedanken).

[a] Gödel writes ‘überall definiertes’ as an insertion after ‘Verfahren’.

Fig. 1.2 Gödel’s Paradox, intuitionistic (Gödel Papers, 12/52, collective item 060772). The other side is given in Fig. 1.1. (Transcription from the Gabelsberger MvA, advised by Robin Rollinger.)

theoretic version of Russell’s Paradox. Wang’s rendition of the classical version (in 8.6.25) follows that in Fig. 1.1 very closely; Gödel evidently had the latter at hand.[79] Having shown it, Gödel made some points that also apply to the intuitionistic version:

(1.82) 8.6.26 The derivation above has no need even of the propositional calculus. Definition by cases is available in Church’s system.[80] It is easy to find functions which are everywhere defined. Unlike the classical paradox,[81] there is no need to assume initially that the crucial concept (or function) of not applying to itself is everywhere defined. The paradox is brief, and brevity makes things more precise. By a slight modification, using provability, it can be made into an intuitionistic paradox. (Wang, 1996, p. 279).

Apparently Gödel did not go on to elaborate that last point, and Wang seems not to have asked.

[79] Gödel wrote on the envelope (in shorthand): ‘erläutert’.
[80] [Note MvA] As shown in Kleene (1934).
[81] [Footnote MvA] This evidently refers to what Gödel called ‘the familiar paradox’ above.

To see Gödel's Paradox (intuitionistic version) as an application of the Fixed Point Theorem, first note the immediate correspondence between the functions in the former and in the latter:

$$\begin{aligned} . &= g \\ E &= h \end{aligned} \tag{1.83}$$

(where the equality is definitional). A is the constructive universe, and the chosen representation principle is formulated in Gödel's footnote: for every totally defined operation (performable series of acts) there is a function (which in the argument is considered as a mathematical object), because these are identified. Note that the case distinction in the definition of the application function is decidable, and that use is made of reflection: if B is a proof that g is everywhere defined, then g is everywhere defined, therefore it can be applied to x. The composition of Gödel's $G = E(x \,.\, x)$ is that of the morphism k, which has a representation as an object in A, namely $a = H = < B, G >$. The definition of a is critically impredicative: it is an object in A, the constructive universe, defined via quantification over A, but there is no construction method yielding all elements of A.

Although for Gödel it was important to comment (because for him it shows something about the depth of the paradox – see section 1.4.5 below) that no propositional logic comes in, for our present purpose it should be observed that its conclusion is readily presented in propositional form:

$$H \,.\, H = 0 \leftrightarrow (H \,.\, H = 0 \to \bot)\,. \tag{1.84}$$

Gödel expresses the same problem in a different way. On the one hand, to the operations defined in Fig. 1.2 the criterion for informal constructivity applies that he had formulated in *Dialectica*: 'die Ausführbarkeit der Operationen unmittelbar aus der Kette der Definitionen ersichtlich' (Gödel, 1958, p. 283n5). On the other hand, as he observes in another note in which he gives the same argument in a slightly different notation:[82]

> Also der Versuch der Wertung von $H[H]$ führt auf einen unendlichen *Regress* (*Church*). Also der Beweis b $\underline{\underline{\text{falsch}}}$. (Gödel Papers, 12/52, collective item 060772) (1.85)

The regress arises because an attempt to evaluate $H \,.\, H$, demands an evaluation of $E(H \,.\, H)$; but the latter demands an evaluation of $H \,.\, H$. In Church's terms referred to above (p. 36), $H \,.\, H$ has no normal form.[83] Intuitionistically, the appearance of this

[82] Gödel now indicates the use of the special application function with straight brackets, and the proof with a lowercase letter.

[83] The role of infinite regress in paradoxes of the Russellian kind was emphasised in Behmann (1931, pp. 41, 42) and Behmann (1959, p. 112). Bernays drew Gödel's attention to (a different aspect of) the latter paper in a letter of October 12, 1961 (Gödel, 2003a, p. 197). But that paper seems to have played no role in Gödel's thinking, or in his exchanges with Kreisel and Wang. It was reviewed in *Zentralblatt* by Ackermann; his criticism was the same as Gödel had made, in correspondence with Behmann, on Behmann's earlier attempts in this direction: no formalism

regress means that there is a problem with the proof of the totality of H, which must be supposed to have shown that a value can be constructed for each argument. A condition of possibility for this problem to arise is the critical impredicativity of the definition of H. In fact, the same regress can be found in the other proof paradoxes; for now, the propositional presentation suffices, and I will postpone a further remark on this to p. 61 in the section on critical impredicativity.

I read the following remark of Gödel to Wang as a comment motivated by his intuitionistic paradox and the other proof paradoxes discussed here.[84]

(1.86) 6.1.13 The concept of concept and the concept of absolute proof [briefly, AP] may be mutually definable.[85] *What is evident about AP leads to contradictions which are not much different from Russell's paradox. Intuitionism is inconsistent if one adds AP to it.* AP may be an idea [in the Kantian sense]: *but as soon as one can state and prove things in a systematic way*, we no longer have an idea [but have then a concept]. It is not satisfactory to concede [before further investigation] that AP or the general concept of concept is an idea. The paradoxes involving AP are intensional – not semantic – paradoxes. I have discussed AP in my Princeton bicentennial lecture. (Wang, 1996, p. 188, 6.1.13, amendations Wang's, italics mine)

AP is the concept of proof independent of any particular formal language or system. It would therefore seem natural to add the concept AP to intuitionism, as intuitionism has always explained truth in terms of such a concept of proof.[86] (It is rather Gödel's proposal in the Princeton lecture to introduce it also in *classical* mathematics that is innovative.) But, Gödel claims, this leads to inconsistency. The similarity between the phrases in (1.86) that I have emphasised and Kreisel's earlier (1.4a) leaps to the eye.[87] Finally, when in (1.86) referring to 'contradictions which are not so much different from Russell's Paradox', Gödel did so while knowing Kreisel's Paradox

is developed and shown to be free of paradoxes (Gödel 2003a, p. 34, letter of April 22, 1931; Ackermann n.d.). Incidentally, Behmann took note of the similarity between Church's ideas on normal form and his own on meaningful predication (1959, p. 122n31).

[84] Note that Wang's (re)presentation of their conversations does not contain an explicit connection between (1.82) and (1.86).

[85] [Note MvA] When Gödel says that the concept of concept and the concept of absolute proof may be mutually definable, he is suggesting, in one direction, an inferentialist theory of concepts, but also, in the other direction, what may be called a conceptualist theory of inference. In particular, the concept of concept would be that which one understands once one understands the inferences that are correct for any concept. For Brouwer, any such *systematic* understanding would intrinsically be a form of applied mathematics.

[86] Since it is the role of AP in mathematics that we are here interested in, Gödel's question whether AP can be treated even independently of any specific system of things (Gödel, 1946, pp. 152-153) is left aside. It is discussed in Crocco (2019).

[87] There is independent reason to believe that Gödel had seen the latter before making this remark to Wang in 1972. Kreisel had the habit of sending his work to Gödel, but in this case there was no need to. By the time of Kreisel's writing (1.4a), Gödel had, like him, become a Foreign Member of the Royal Society (Gödel in 1968, Kreisel in 1966) and on December 17, 1969, Kreisel wrote in a letter: 'Wahrscheinlich kriegen Sie bald von der Royal Society die 1969 Obituary Memoirs, einschließlich den Nachruf auf Brouwer, den ich gemeinsam mit dem Topologen M. H. A. Newman

and Goodman's Paradox, and, very likely, also Troelstra's Paradox, as seen in the respective sections here; all of which are like Russell's Paradox in the sense explained in section 1.4.1.[88]

There are two circumstances that I find suggestive of the idea that, moreover, Gödel's (intuitionistic) Paradox, at least as presented in Fig. 1.2, was occasioned specifically by Kreisel's Paradox. First, Gödel's footnote addresses (and denies) exactly the distinction that Kreisel's letter to Gödel quoted in (1.79) turns on. Second, Gödel's Paradox illustrates the following theme in Kreisel's paper even more explicitly than Kreisel's:

> Finally, *isolation of primitive concepts*, in terms of which the other can be defined, and laws (axioms) for these primitives. Current candidates are *construction* (function) and the *application operator* with *proof* as a suppressed parameter. (1.88)

which comes with the footnote

> As *ordinal* and *order* of the cumulative type theory are suppressed in the practice of set theory. The occurrence of such hidden parameters seems essential in work that gives an *analysis* of informal mathematics. (Kreisel, 1965, 2.1, p. 121, original italics) (1.89)

It is certainly essential to intuitionistic mathematics, which by its nature is informal, that functions are applied only to objects that have (actually, or hypothetically) been proved to be in their domain.

In the other direction, even on the (as yet unsupported) supposition that Kreisel knew (the content of) Gödel's Paradox when he set out to write his paper for the Saaty volume, it is clear that he would have had little direct motivation to include it: his interest there is in (further) developing an alternative interpretation of formal intuitionistic logic, not in *using* the intuitionists' own understanding. It is only when introducing the convention in 2.151 for use in his theory of constructions that he makes a short detour to show his paradox. Similarly, the topic of Kreisel's 1971 publication in which he presents the classical version of Gödel's Paradox, generalisations of recursion theory, would hardly have motivated including Gödel's intuitionistic Paradox or, for that matter, his own.[89]

---

verfaßt habe.' (Kreisel Papers, 50/2). The index Dawson (1984) shows on p. 68 that Gödel owned the Society's *Biographical Memoirs* for 1968–1977.

[88] Of the earlier (voluntarily brief) discussion of (1.86), that in Crocco (2019, p. 571), I note that it does not give an example of a contradiction that AP leads to, except in its

> Remark 2. According to Gödel, intuitionism rejects the use of unrestricted universal quantifications (all objects, all proofs, etc.) and therefore extensional and intensional paradoxes do not appear in it. Absolute provability implies reference to all proofs that can be performed by a human agent in any domain. In this sense it is in contradiction with intuitionism. (1.87)

But that contradiction is, in the way in which it arises, not similar to Russell's Paradox. Furthermore, the first half of the first sentence is problematic, both in the claim ascribed and in the ascription (following Wang) of that claim to Gödel; see the discussion of (1.124) and (1.125) below.

[89] In fact, there Kreisel (1965) is refered to only once (p. 166), and on another topic.

As stated at the end of the previous section (p. 38), it is highly likely that Gödel and Kreisel discussed the diagnosis in Kreisel's letter of April 1, 1968, and may have done so on Kreisel's travel to the Buffalo conference, but I have no direct evidence for either supposition. However, in Kreisel's first published remarks on notions and functions after that letter – the published version of his address at Buffalo – he abandons the categorical distinction he had appealed to between *Verfahren* and *Funktion (Konstruktion)*, where the latter come with a domain of definition but the former do not. Thus he comes closer to Gödel's view in the footnote in Fig. 1.2, and I take this to be indirect evidence of a discussion with Gödel. Kreisel's view now is as follows.

(1.90) Let us scrutinize a bit the basic relation:

> For some given notion α, the construction (more precisely, judgement) c proves αd for variable d.
>
> [...]
>
> If we think of the variable d as ranging over the, so to speak, absolutely unattained universe of all constructions, it seems dubious that there should be any construction (something that we grasp completely) which proves αd, even if we have convinced ourselves that, for any clearly given d, α is indeed a well defined notion.
>
> [...]
>
> On the other hand, if we take some particularly simple notion αd, say $\beta d \triangleright \beta d$ (where I use $\triangleright$ for truth functional implication [...]) we simply have a proof. Whatever else may be in doubt, we have a perfectly clear idea or 'schema' for verifying $\beta d \triangleright \beta d$. The kind of judgement involved here plays the same role among proofs as, say, the identity operator plays among functions. It is simply a mindless ritual to chant: for each type we have a different identity operator. (Though, trivially, for each domain D the set of pairs $\{<x, x>: x \in D\}$ depends on D.)
>
> The obvious and immediate conclusion is: just as there are some operations which are defined for arbitrary operations (in the non-trivial sense of giving distinct values for 'lots' of arguments, e.g. the identity operator, the composition operator etc.) so there are some notions α which can be proved by constructions to hold for unrestricted d. The definition of other operators depends essentially on a given domain ('essentially' in the sense that the function is made total by a trick of, say, defining its value to be zero outside the given domain; cf. [21], 2.151, pp. 124-125).[90] In the case of notions α, the corresponding restriction concerns the variable d. (Kreisel, 1970a, pp. 129–130)

This entails that a diagnosis of his paradox must, unlike the one in (1.79), depend on something different than a complete separation of notions and functions (constructions). In this respect, it is telling that Kreisel, when the next year he published the classical version of *Gödel*'s Paradox (without the name), he remarks on the problem of evaluating $H\ .\ H$ (notated differently), as Gödel does in (1.85) above. Kreisel

[90] [Note MvA] The remark referred to is contained in quotation (1.73) above.

proposes to 'look at the steps of the argument by the light of nature', which means, in particular, that application of a function rule presupposes that a value has been assigned to its argument; but, to borrow his words, one is not given even the remotest hint of how that value is to be determined (1971, p. 191). Exactly the same could (and, because of its constructive context, should) be said about his own paradox. However, there was *no* corresponding remark either in Kreisel (1965) or in Kreisel's letter of 1969, as quoted above in (1.73) and (1.79).

Bew⟨eis⟩-Begriff im Intuit⟨ionismus⟩.[a]
(Zeitabhängigkeit)

Df Eine Funktion ist etwas, von dem erkannt wurde, dass es immer einen definierten Wert
$\mathfrak{A}_t(f)$ $\quad$ $f(x)_t$
hat wenn ein Argument gegeben ist (aber wenn darin ⟨das⟩ Argument zeitlich gegeben ist, kann sie einen anderen Wert ⟨annehmen⟩)

$t = 0$ *Th* $F(x) = 0, F(0) = 1$ ist eine Funktion $\quad \mathfrak{A}_{0t}(F) \quad x =_t y \equiv x =_{t_1} y$
1 *Th.* $(f \,.\, x)_t = 0$ wenn ein $\mathfrak{A}_{0t}(f)$ ⟨zeitabhängig ist⟩
$=$ Wert wenn $\mathfrak{A}_{0t}(f)$ zeitunabhängig
. ist eine zeitabhängige Funktion für $t > 1$
2 *Th* $F(f \,.\, f_t)$ ist eine Funktion G für $t > 2$
$G \,.\, f_t = F(f \,.\, f)_t$ gilt jederzeit
Th $G \,.\, G = F(G \,.\, G)$

Zeitabhängigkeit von $x \,.\, y$ (= Anwendung)
(andere Methode der „Einsicht" beziehungsweise des Errechnens)

[a] This title and the parenthesis are the text on one side of the paper, the rest that on the other.

Fig. 1.3 Gödel's Paradox, time-dependent (Gödel Papers, 12/52, collective item 060772). (Transcription from the Gabelsberger MvA, advised by Robin Rollinger.)

To return to Gödel's intuitionistic Paradox, note the variant in Fig. 1.3. In the definition of F, the condition has to be understood as that on E in Fig. 1.1. The value of the function f applied to x at time t is notated $f(x)_t$, and $\mathfrak{A}_{0t}(f)$ is the sequence of the arguments to which f has been applied between times 0 and t, such as they were given to us at these moments (i.e., intensionally), and possibly with repetitions. I take it that Gödel labels the definitions of the functions $F, . ,$ and G as theorems because at the same time the existence of these functions is established, a construal of definitions similar to that in his introduction of 'reductive proofs' in the revised Dialectica paper (Gödel, 1972, p. 275, note h1). These theorems are themselves established in time.

Note that the case distinction in the definition of the application function is decidable, with the particularity that which of the two conditions is proved to hold may, for the same arguments f and x, change with time. An example of time-dependency of $f \,.\, f$ would be an identity function f, calculated at t_{37} by the projection $f(f) = \pi_0 < f, f >$ and at t_{40} as $f(f) = f$ ('andere Methode'); then for $t > 40$, $(f \,.\, f)_t = 0$. We will see something very similar in Troelstra's Paradox.

Gödel offered no solutions to his paradoxes; I will come back to that at the end of section 1.4.5.

1.4.4.3 Troelstra's Paradox

Kreisel intended to enrich his Theory of Constructions with axioms for the intuitionistic 'thinking subject', with an eye on reconstructing Brouwer's so-called 'Creating Subject' arguments (Kreisel, 1967a, p. 180).[91] Those arguments are another example of (making explicit, or exploiting) a hidden parameter; see quotation (1.88) above. There the hidden parameter in question was proof, here time. (There is no reason these could not be treated together, but there seems to be no experience with that.)

Troelstra added an axiom to Kreisel's two, and in a modern formulation, the axioms and their intended meanings are as follows.

$$\forall n(\Box_n A \vee \neg\Box_n A) \tag{CS1}$$

That is, for any stage, it is decidable for the Creating Subject whether by that stage it has made A evident.

$$\forall n \forall m(\Box_n A \to \Box_{n+m} A) \tag{CS2}$$

The Creating Subject never forgets what it has made evident.

$$\exists n \Box_n A \leftrightarrow A \tag{CS3}$$

A proposition A is true if and only if the Creating Subject has made A evident by some stage. From left to right this is a reflection principle: there is a certain proof for A, therefore A.

These axioms show that about provability in Brouwer's sense 'one can state and prove things in a systematic way', as Gödel might have said it (see quotation (1.86)).

But Troelstra had also found a paradox. He discussed it with Kreisel at the Buffalo conference in 1968;[92] the time span in which Kreisel was preparing his part of the Brouwer obituary included all of that year (see the end of footnote 42 above). It is treated in the notes of the lecture series Troelstra gave there, his influential *Principles of Intuitionism* (1969), of which, incidentally, Gödel owned a copy (Dawson, 1984).[93] Troelstra published on it again in Troelstra and van Dalen (1988, ch. 16, section 3) and Troelstra (2018).

[91] Kreisel had received a letter from Kripke in which the latter had proposed a weak version of the Brouwer-Kripke Schema. See van Atten (2018, p. 1588).

[92] Email Troelstra to MvA, May 1, 2016. Troelstra spent the academic year November 1966-November 1967 with Kreisel at Stanford. Kreisel rarely refers to Troelstra's Paradox in print; a place where he does is Kreisel (1972, p. 326).

[93] Dawson's list shows that he also had *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis* (Troelstra, 1973a).

For the present purpose, the last publication is the clearest and most useful one. It is best read as an argument in which we put ourselves in the shining shoes of the (idealised) Creating Subject, which thus reasons about itself:[94]

> Let us now use α, β to denote arbitrary, not necessarily predeterminate, and not necessarily infinite sequences of natural numbers, and let us consider statements of the form 'α is a total sequence'. For example, if α is defined as a primitive recursive sequence, this conclusion is immediate as soon as α is defined. If α is initially given to us as a partial recursive function, we may at a later stage conclude that α is a total sequence, namely if we have found a proof of this fact. A lawless sequence is from the moment it is initiated a total sequence. (1.91a)
>
> The original idea for the paradox was as follows. Let α^n be the n-th total sequence the C[reating] M[athematician] encounters when running through the stages of activity; then consider a sequence β defined by
>
> $$\beta(n) = \alpha^n(n) + 1$$
>
> β is total, and at some stage m β should appear as an α^n. But then $\beta(n) = \alpha^n(n) = \alpha^n(n) + 1$, a contradiction. This is just a classical diagonalization argument.

Self-application of functions is not a feature of Troelstra's background theory, but its effect is provided for by letting the natural numbers also play the role of indices (a form of coding) to the encountered sequences $\mathbb{N} \to \mathbb{N}$ (functions). In this argument, one finds the principles CS1–3 instantiated as follows. CS1: Such an encounter consists in the act of proving that the sequence is totally defined, and for the Creating Mathematician it is decidable whether at a given stage it has such an encounter. CS2: A necessary condition for a list of the α^n is that these encounters are not forgotten. CS3 (from left to right): If, at some stage, it is proved that a sequence is totally defined, then that sequence is totally defined. CS3 (from right to left): If it is true that β is totally defined, then this is proved at some stage m.

> Mark van Atten observed that perhaps β is not well-defined, because, having encountered α^n, we are not certain how long we have to wait before the next total sequence appears. (1.91b)

That is, in its original formulation Troelstra's Paradox depends on an appeal to an unacceptable version of Markov's Principle: for discussion, see van Atten (2017c).

> This can be remedied as follows. At stage 0 we take α^0 to be the constant zero function. As long as no new total sequence is declared at stage $n + 1$, we take α^{n+1} to be equal to α^n; and if at stage $n + 1$ a new total sequence γ is found, we take α^{n+1} to be equal to γ. Then we can diagonalize as before. (Troelstra, 2018, p. 14) (1.91c)

[94] See van Atten (2018) for further discussion and references.

In terms of the Fixed Point Theorem, the representation principle in this Paradox is the instantiation of CS3: every totally defined sequence (function) $\mathbb{N} \to \mathbb{N}$ that the Creating Mathematician encounters can be correlated to an element of $\mathbb{N}$.

The application function g is defined as

$$\begin{aligned} g : \mathbb{N} \times \mathbb{N} &\to \mathbb{N} \\ < n, x > &\mapsto \alpha^n(x) \end{aligned} \tag{1.92}$$

Since the difference between the two versions that Troelstra described lies only in the way that the list of the α^n is constructed, g is the same for both. Note that the constructivity of g depends, via its dependence on that list, on the decidability of the appearance, at a given stage, of a total sequence.

The role of the morphism h is here played by the successor function. The morphism $k : \mathbb{N} \to \mathbb{N}$ yielded by the Fixed Point Theorem is a recipe for assigning a natural number to each natural number, and thus totally defines a sequence, Troelstra's β. The Creating Mathematician has this insight at some stage of its activity (otherwise it would, intuitionistically, not be true), which means that $\beta = \alpha^n$ for some n. Thus, n represents β with respect to g.

We now have the contradiction $\beta(n) = g(n, n) = h(g(n, n)) = g(n, n) + 1$. To relate this to our theme in propositional logic, change h to the two-valued

$$\begin{aligned} h : \mathbb{N} &\to \mathbb{N} \\ x &\mapsto \begin{cases} 0 & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases} \end{aligned} \tag{1.93}$$

This yields $g(n, n) = 0 \leftrightarrow g(n, n) = 1$, that is,

$$g(n, n) = 0 \leftrightarrow (g(n, n) = 0 \to \bot)\ . \tag{1.94}$$

The definition of the $n \in \mathbb{N}$ that represents β is impredicative, as it lies in the range of the latter's argument. It is furthermore critically impredicative, because this n, and any n in an argument $< n, x >$ of g, is viewed here not as a natural number as such, but as a natural number in the role of an index into the collection of the encountered total sequences, and in that sense as dependent on the latter. But there is no construction method for that collection, as the Creating Subject is free to go about its constructive activity as it pleases. This critical impredicativity exists on both the original and the remedied formulations of Troelstra's Paradox.

The paradox can be seen as a special case of the intuitionistic version of Gödel's Paradox in Fig. 1.2, if one accepts the idea that the definition of $f = < B, g >$ there may change in the sense that B initially is, say, the object 0, but is identified with the proof that the operation g always yields a result as soon as there is one.[95] If the

---

[95] Troelstra writes:

(1.95) Originally, I used, instead of 'total sequence' the notion 'a total sequence determined by a recipe'. I used the word 'recipe' instead of 'lawlike', because I did not want to suggest that the sequence was recursive, only that it was fixed by a recipe relative to the activity relation to time in such a change is made explicit (and we furthermore allow for partial functions), we get a version of Gödel's time-dependent paradox in Fig. 1.3, with the application function

$$(f\ .\ x)_t = \begin{cases} 0 & \text{if by stage } t,\ f \text{ has not been proved to be total} \\ f(x) & \text{if it has} \end{cases} \tag{1.96}$$

This definition by cases is governed by a decidable disjunction, with the property that it is time-dependent which of the two disjuncts is provable. In both Gödelian renderings of Troelstra's Paradox, the indirect self-reference of Troelstra's β via its index number is made direct in the application of a function to itself. As alluded to at the beginning of this section, it can, at present, not be excluded that Gödel devised his paradox after Troelstra and had seen the latter's, but I have neither positive nor negative evidence for that.

One thematic solution that Troelstra proposed was to stratify:

> To each mathematical assertions and construction we suppose a level (of self-reflection) to be assigned. [. . .] Assertions which may be understood or constructions which can be carried out without reference to $\vdash_n$ $[\Box_n A]$ are said to belong to level zero. (1.97)
>
> Assertions which are described using $\vdash_n$ A for A of level p and constructions of level p, are said to belong to level $p + 1$. Likewise, constructions defined relative to $\vdash_n$ A for A of level p are said to be of level $p + 1$.
>
> [. . .]
>
> [O]ur paradox cannot be derived anymore.

Indeed, the critically impredicative definition of β is ruled out, as its construction can now be carried out only at a higher level than that of any collection of sequences that its definition can legitimately refer to. It is noteworthy that Troelstra proposed it only as an 'approach [that] deserves further investigation' (Troelstra, 1969, p. 107), and when writing about it again in *Constructivism in Mathematics*, he qualified it as 'at least as problematic' as the theory it replaces, unfortunately without expanding (Troelstra & van Dalen, 1988, p. 846). In Troelstra (2018), on the other hand, he describes it neutrally.

> of the CM in general. But in view of the fact that the CM is completely free in his actions, a 'sequence determined by a recipe' can be as un-predetermined as an arbitrary choice sequence.

In fact, in the original publication the term had been 'lawlike' (Troelstra, 1969, p. 105); it was changed to 'fixed by a recipe' in (Troelstra & van Dalen, 1988, p. 845). Be that as it may, the subsequent crystalisation into 'total sequence' in the 2018 version has the effect of bringing it even closer to Gödel's Paradox.

1.4.4.4 Goodman's Paradox

Goodman's Paradox (which is what is often really meant when the 'Kreisel-Goodman paradox' is referred to) first appears in the introduction to Goodman's dissertation. He introduces (p. 4) an operation π, assumed (with Kreisel) to be decidable, such that

$$\pi(g, y) \leftrightarrow y \text{ is a proof of } \forall x(g(x) = 0)\,, \tag{1.98}$$

and combines this with Kreisel's understanding of intuitionistic implication in (1.2) above to arrive at a contradiction:

(1.99) To recapitulate briefly, we have said that a pair (p, f) is a proof of the proposition $A \rightarrow B$ just in case p is a proof that, if q is any proof of A, then $f(q)$ is a proof of B. So far we have no way of excluding the possibility that q is itself built up in some way from p. It is largely this impredicative character of implication that makes the theory of constructions interesting from a technical point of view. Indeed, the most natural formalization of the conception we have outlined so far is inconsistent. It suffices to construct, using π, a function f such that $f(x) = 0$ if and only if $x(x)$ is a proof that no y proves that $f(x) = 0$. Now suppose that y proves that $f(x) = 0$. Then $f(x) = 0$, and so no y proves that $f(x) = 0$. This contradiction, together with the decidability of the proof predicate, shows that no y can prove that $f(x) = 0$. Therefore there must be a function g such that, for any x, $g(x)$ proves that no y proves that $f(x) = 0$. In particular, $g(g)$ proves that no y proves that $f(g) = 0$. That is, $f(g) = 0$. Hence there is a proof that $f(g) = 0$, which is absurd.

Goodman does not spell out how f is constructed, but the formal derivation of this paradox he went on to give in Goodman (1970),[96] which closely follows this informal one, naturally suggests that it is a fixed point construction. It is that formal derivation that, read according to its intended interpretation, will be reconstructed here, in different terms.

For the sake of presentation, certain simplifications are made to abstract from some of the low-level machinery in Goodman's own setting, which is an extended type-free λ-calculus. That machinery was there of course to serve the purpose of his project, a theory of constructions as a foundation for logic and arithmetic, not for its ease of use in generating a paradox; and it remained in place after Goodman's repair (see the end of this section).[97] I presuppose that an arithmetical language has been

---

[96] Both Gödel and Heyting read that paper closely. Gödel's notes to that paper (and others in the same volume) are in Gödel Papers, 10a/40, collective item 050142 Heyting summarised, in *Mathematical Reviews*, Goodman's paper with a fair amount of detail, including the impredicativity and Goodman's solution; but, unfortunately, without comments (Heyting, n.d.).

[97] The mended theory did not prove viable; see Kreisel (n.d.-a, n.d.-b, n.d.-e); Weinstein (1983, p. 265); Dean and Kurokawa (2016, pp. 53–54); van Atten (2017b, section 4). Note that Goodman came to hold a view that is critical of his own efforts in a different and farther-going way:

fixed and that sufficiently much arithmetic has been developed to implement Gödel-numbering; and I assume that we have an implicational logic for decidable statements from the outset, whereas Goodman constructs it first. Where the reconstruction does not simplify is in its attempt to bring out the exact reasoning in Goodman's Paradox. (The reader should consult also the rich discussion by Dean and Kurokawa (2016), and compare their reconstruction and mine.[98] They use the combinator Y and their perspective is that of a comparison with (and to) Montague's Paradox. For the reason given on p. 29 above, I prefer to use the Fixed Point Theorem, and, as per section 1.4.1, my emphasis is rather on the paradox's relation to the propositional reasoning pattern that Kreisel mentions in (1.4c).)

Let A be the universe of constructions, and B the set of truth-values $\{\top, \bot\}$, considered as two arbitrary but distinct constructions. The representation principle here is: every morphism from A to B is also a construction (an element of A).

Let P be the decidable binary proof predicate, defined on constructions, 'v is a construction that proves the proposition whose statement has Gödel-number u'.[99]

Let $f(v, u) \colon A \times A \to B$ be the characteristic function of P, and let the family $f_v \colon A \to B$ be given by $f_v(u) = f(v, u)$. Part of the assumed decidability of P is the assumption that P is everywhere defined; if it is, then so are the functions f and f_v.

Define an application function for the f_v by

> I myself have been attracted by intuitionism. But I have gradually come to see that, in the long term, strong intuitionistic convictions undermine one's actually doing mathematics. By embracing intuitionism the mathematician is giving up the most powerful motivation for his work — the search for publicly validated truth. [...] There is a sense in which intuitionism is inadequate in its own terms, for it overlooks what is introspectively obvious: that I am interested in my constructions not for their own sake but for the new truths they enable me to find. [...] Just as the constructions lie behind the symbols and give them their interest and meaning, so there is something behind the constructions – mathematical truth. (1.100)
>
> [...]
>
> Mathematical truth, unlike a mathematical construction, is not something I can hope to find by introspection. (Goodman, 1979, p. 545)

Goodman's notion of truth is different from Brouwer's. For a defense of intuitionism against the charge of solipsism, see Placek (1999) and van Atten (2004, ch. 6).

[98] In the derivation in Dean and Kurokawa (2016, pp. 43–44), the definiens of $h(y, x)$ should be replaced with $\lambda y.(\pi y x \supset \bot)$ and the fixed point equation with $Y(h(y, x)) \equiv h(y, x)(Y(h(y, x)))$. Here $\supset$ denotes a term representing truth functional implication whose details will depend on which terms are chosen to represent $\top$ and $\bot$ in the untyped lambda calculus (Goodman 1970, p. 105, and continued in Barendregt 1984, p. 44) but is otherwise similar to Goodman's definition of $\supset_k$ (Goodman, 1970, p. 106). I thank Dean and Kurokawa for correspondence which led to these points. Since the approach and structure of their reconstruction is left unaltered, these changes make no difference for the discussion of the philosophical questions involved, be it in the reconstruction itself, or in comparison with other reconstructions. On such a comparison, see footnote 102 below.

[99] Goodman, in effect, defines $P(v, u)$ as 'v is a proof that $u(z) \equiv \top$ for all z'.

$$
\begin{aligned}
g\colon A \times A &\to B^A \\
< j, x > &\mapsto \begin{cases} j(x) & \text{if } j = f_w \text{, for some } w \\ \lambda z.\bot & \text{otherwise} \end{cases}
\end{aligned}
\tag{1.101}
$$

The equality here is intensional. Note that g is everywhere defined because the case distinction is decidable and the f_w are everywhere defined.

Define the family of morphisms

$$
\begin{aligned}
h_v\colon B^A &\to B^A \\
\lambda z.f(v,z) &\mapsto \lambda z.f(v(v), \ulcorner P(v,z) \to \bot \urcorner) && \text{if } v(v) \text{ is defined} \\
\lambda z.t[z] &\mapsto \lambda z.\top && \text{in all other cases}
\end{aligned}
\tag{1.102}
$$

Constructivity of the image for all arguments presupposes that, for the chosen value of the parameter v, it is decidable whether $v(v)$ is defined, and that, if it is defined, f is constructive (i.e., that P is decidable). Under those presuppositions, h_v is everywhere defined; I will take the presupposition concerning f as a given, and to find a value for v is precisely how the argument will proceed.

Applying the Fixed Point Theorem to g and h_v, we first obtain the morphism $k_v\colon A \to B^A$; by the chosen representation principle, it is represented by an element of A, which is k_v itself. Then we get the fixed point $b_v\colon A \to B = k_v k_v$, and have $b_v = g(k_v, k_v) = h_v(b_v)$. Hence for all z, the truth values given by $b_v(z)$ and $(h_v(b_v))(z)$ are identical.[100]

In propositional terms, this entails that

$$
P(v,z) \leftrightarrow P(v(v), \ulcorner P(v,z) \to \bot \urcorner) \, . \tag{1.103}
$$

On account of the decidability of P, here and in the remainder of this section, the implication can be interpreted truth-functionally (which for the reductive purpose of the Theory of Constructions would be required).

As an instance of reflection ('What is proved, is true'),

$$
P(v(v), \ulcorner P(v,z) \to \bot \urcorner) \to (P(v,z) \to \bot) \, , \tag{1.104}
$$

so with (1.103)

$$
P(v,z) \to (P(v,z) \to \bot) \, , \tag{1.105}
$$

and by contraction as in (1.53)

[100] The family h_v corresponds to Goodman's term $\mathfrak{h}$ (1970, p. 108), the fixed point b_v to his $\mathfrak{a}$, and the equality $b_v(z) = (h_v(b_v))(z)$ to his equivalence $\mathfrak{a}z \equiv \mathfrak{h}\mathfrak{a}z$. Goodman does not explain how he arrives at his fixed point; one notes that the term $\mathfrak{a}$ can be obtained by applying the combinator Y to $\lambda y.\mathfrak{h}yz$ and abstracting on z. Also, in his explanation on p. 103 why he designs a theory of partial functions, the term that he shows to be undefined is in effect an application of Y to the term $\mathfrak{f}$ as he defines it there. It seems safe to say that Goodman is there adapting Kreisel's Paradox.

$$P(v, z) \to \bot\,. \tag{1.106}$$

Thus, using the Fixed Point Theorem we have obtained a morphism c that maps any construction v to a proof of $P(v, z) \to \bot$ for all z, provided that $v(v)$ is defined.

Again by the representation principle, we have $c \in A$, so $c(c)$ is defined, whence it is admissible to choose c as value for the parameter v. If furthermore we choose an arbitrary z_0 for z and repeat the above schematic reasoning for those values,[101] we conclude that the construction $c(c)$ is a proof of $P(c, z_0) \to \bot$, and hence

$$P(c(c), \ulcorner P(c, z_0) \to \bot \urcorner)\,. \tag{1.109}$$

Along the way to (1.109), we obtained instances of (1.103)–(1.106), in particular

$$P(c, z_0) \leftrightarrow P(c(c), \ulcorner P(c, z_0) \to \bot \urcorner) \tag{1.103$'$}$$

and

$$P(c, z_0) \to \bot\,. \tag{1.106$'$}$$

Combined with (1.109) itself, these yield

$$P(c, z_0) \tag{1.110}$$

and then

$$\bot\,. \tag{1.111}$$

[101] On the (constructive) relation between the proof of a general statement and proofs of its instances, there is Weyl's well known conception of the former as an 'Urteilsanweisung' (Weyl, 1921). I should like to recall here also Freudenthal, who wrote:

> der einmal gelieferte allgemeine Beweis dient uns nicht mehr als eine Landkarte, die uns die Bergbesteigung zwar erleichtert, aber nicht erspart (Freudenthal, 1936, p. 116), (1.107)

with the footnote

> Man könnte meinen, daß dies Immer-wieder-von-neuem-Beweisen nicht nötig ist bei Hilfssätzen, die sich als explizite Formel darstellen, wie $m + n = n + m$. In Wirklichkeit bleibt einem aber weiter nichts übrig, als die Umordnung, von der diese Formel handelt, immer wieder, wenn sie nötig ist, von neuem vorzunehmen. Natürlich wird man in der sprachlichen Darstellung des Beweises das nicht tun, aber das sagt nichts gegen unsere Feststellung und alles gegen die sprachliche Darstellung. (1.108)

Note that $P(c, z_0)$ here is derived via a proof of $P(c, z_0) \to \bot$,[102] which illustrates one half of Kreisel's remark (1.4c); but as no proof of $(P(c, z_0) \to \bot) \to P(c, z_0)$ is involved, it does not illustrate that remark fully. This it has, mutatis mutandis, in common with the reasoning in Myhill's Paradox, as reconstructed in (1.61)–(1.63). As remarked there, that reasoning cannot be restructured to the desired effect. In contrast, in the present case, where the constructive logic of $\to$ is truth-functional, from (1.110) we could conclude to $(P(c, z_0) \to \bot) \to P(c, z_0)$, and proceed from there. But that restructured proof, deriving its conclusion twice, would of course not be natural (see also the remark following (1.56)).

The definition of the representation k_c of the morphism k_c is critically impredicative, because it depends on a quantification over A, for which there is no construction method. Goodman's solution of this paradox was to stratify the universe of constructions, which, in effect, invalidates the definition of the representation. (By footnote 102 above, the effect he intended was a different one; he did the right thing for the wrong reason.) His criterion for the stratification is 'the subject matter of proofs'. That is a much broader criterion than Troelstra's of levels of self-reflection (section 1.4.4.3), and correspondingly more difficult to justify (see the discussions referred to in footnote 97 above).

1.4.5 Critical impredicativity

A correct but limited answer to the question why these proof paradoxes arise would be to say that their contexts satisfy the hypothesis of the Fixed Point Theorem. To see if a more specific cause can be identified, consider the following common aspects of the applications of that theorem in these paradoxes:[103]

---

[102] This is the parallel to the fact in Goodman's own version that $\vdash \pi(\mathfrak{gaf})\mathfrak{b} \equiv \top$ is arrived at via $\vdash \pi(\pi(\mathfrak{gaf}) \supset_1 (K\bot))(\mathfrak{ff}) \equiv \top$, where $\mathfrak{b}$ is obtained as an internalisation of the preceding derivation, and thereby depends on the proof $\mathfrak{ff}$ (1970, pp. 109). Thus, a proof of the antecedent of an implication is given that depends on a proof of that implication. In this sense, Goodman's derivation of a paradox, and the reconstruction in the text, are such that they at least potentially confirm his concern over the impredicativity in his initial definition of implication (1970, p. 109); and Goodman clearly thought that his paradox *does* confirm that concern. In contrast, (*i*) in the text I argue that there is a second impredicativity in play, that of the representation principle, which remains implicit in Goodman but which is constructively not acceptable; (*ii*) if one accepts the argument discussed around (1.115) below, the impredicativity in implication is constructively acceptable; (*iii*) although by (*i*) and (*ii*) the problem that Goodman's Paradox brings to light is different from the one he thought it was, it so happens that the solution that he proposed (stratification) is also a solution to the actual problem. Finally, note that in the derivation in the reconstruction of Goodman's Paradox in Dean and Kurokawa (2016, pp.43–44) there are no analogues to Goodman's z and $\mathfrak{f} - v$ and c in my reconstruction. This absence shows that Goodman's concern over impredicativity of implication that motivated his Paradox plays no role in the paradox they derive. The philosophical significance of this difference would seem to be context-dependent, and is a question I leave aside here.

[103] Compare the analogous section 5 in the discussion of Goodman's Paradox in Dean and Kurokawa (2016).

1. *Informal reflection.* 'There is a contentual proof of p, therefore p.' This is part of the informal constructive explanation of truth.
2. *Informal decidability.* In each of the proof paradoxes as analysed above, the definitions of the morphism g – in (1.74), (1.83), (1.92), and (1.101) – depends, for its constructivity, on a notion in Kreisel's sense. To repeat: such a notion is a property that is decidable, not necessarily in a mechanical way, but for those who understand the terms (Kreisel, 1965, pp. 2.13, 2.141). The same dependence exists for h – in (1.75), (1.83), (1.93), and (1.102) – , where in Troelstra's Paradox (1.93) this is just the decidable equality on the natural numbers, the simplest property of the kind that notions were in fact introduced to generalise (Kreisel, 1965, p. 123, 2.13).
3. *A representation principle.* The motivation for accepting the principles figuring here was given together with (1.65) above.
4. *Critical impredicativity.* The general reason for the appearance of an impredicative definition in each was given in the comments on the Fixed Point Theorem (p. 28, point 4); the reason why they are critical, in the discussion of the respective cases.

I will argue that it is the last aspect that is responsible for these paradoxes. I will not attempt to show that the other three aspects, doing away with any of which would also block these paradoxes, are beyond doubt (although to traditional intuitionism, and also to me, in the present cases, they are). Rather, my take will be that rejecting the fourth is in any case necessary, and clearly is sufficient.

The following discussion is constrained (constricted) by the choice to organise it, like the present paper as a whole, around Kreisel's views – here, in particular remark (1.4b) – and their development. On the one hand, this organisation is natural, to the extent that Kreisel was one of the main participants in the discussion of the relations between constructivity and impredicativity, and an influential one at that, as seen in writings of intuitionists after Brouwer and Heyting such as Troelstra and Van Dalen (but hardly in those of the Nijmegen School).[104] On the other hand, it leaves no natural occasion for reflection on what an analysis of the views of Brouwer, who was neither a participant in that discussion nor among Kreisel's (epistolary) contacts,[105] might suggest about the matter once these are treated in their own right. (See the questions (*i*)–(*iii*) raised, but not taken up, on p. 66 below.) In contrast, ample attention is given to the views of Gödel, which strongly influenced Kreisel's thought in question. But in the end we will rather see the emergence of strong divergences between Kreisel's and Gödel's views.

An early occasion on which Kreisel brings up impredicativity is his 1959 lecture 'La prédicativité' (Kreisel, 1960), but there the emphasis is, as the title suggests, on seeking positive characterisations of predicativity. For the present purpose, more

[104] Wim Veldman informs me that there has been essentially no direct contact between members of the Nijmegen School and Kreisel either.

[105] See footnote 63.

useful is a complementary lecture in 1962 by (his close contact) Bernays,[106] which refers to Kreisel's but makes, as its title says, 'Remarques sur l'imprédicativité' (Bernays, 1962):[107]

(1.112) Une définition d'un objet d'une espèce S (d'un nombre, d'un point) - en bref : « d'un S » est imprédicative s'il intervient une quantification par rapport aux S. C'est l'imprédicativité au sens général.

Le cas critique d'imprédicativité se présente si les conditions suivants sont remplies :

1. Une définition contenant une quantification par rapport aux S est nécessaire pour démontrer l'existence d'un S ayant une certaine propriété.
2. L'espèce S n'est pas celle des individus, mais, pour ainsi dire une espèce dérivée : espèce de fonctions, de suites, de prédicats, de classes.

The first part of the second condition I take to mean that the members of S are neither given to us from the outset, nor generated by a construction method. In the latter cases one speaks of an 'impredicative specification' or 'characterisation' of an object.[108]

In my view, the term 'critical' may be said to apply here in two of its meanings (in English as well as in the French of the original):[109]

1. 'Crucial'. In the critical case, quantification over the domain is our only way to form an intention towards the object defined.[110]
2. 'Of the nature of, or constituting, a crisis'. If the definition of an object is critical in the first sense, then furthermore the existence of that object can be doubted or even rejected to the extent that the correctness of critically impredicative definition itself can.

The standard constructivist objection is that if a definition of an object is critically impredicative, then this definition cannot be used to guide a construction process of that object. For then, in order to construct the object, we would first have to construct the domain quantified over,[111] but as this domain contains the object we are in the

[106] Their contact started with a letter from Kreisel in 1947, and lasted until Bernays' death in 1977. See for some details Isaacson (2020, p. 109).

[107] There is, of course, no novelty in Bernays' characterisation of impredicativity; I use it because it is well-phrased, because I find its terminology useful, and because Bernays' paper is of historical interest. (On the latter aspect, see footnote 113 below.)

[108] E.g., Ramsey's definition of a certain man as 'the tallest in a group' (Ramsey, 1931, p. 41), or 'the smallest natural number n such that $P(n)$.'

[109] As given in, respectively, the *Oxford English Dictionary* and the *Trésor de la langue française informatisé* (http://atilf.atilf.fr/tlf.htm).

[110] A critically impredicative definition need not be unique, but equivalents will likewise depend on such a quantification.

[111] There are, of course, certain quantified propositions that can be given constructive meaning without requiring a construction of the objects in the domain quantified over, namely by an appeal to a continuity principle or to a general property of the objects forming that domain; but such appeals do not serve to describe a unique object in the domain.

process of constructing, we find that we can only complete that process if we already have completed it. In such cases the circularity of the impredicative definition is vicious.[112]

Returning to Bernays' text, we see that further on he comments on the Proof Explanation:

> Mais la question se pose si l'intuitionnisme se restreint à des raisonnements prédicatifs. Je crois que ce n'est pas le cas. En effet, dans les raisonnements intuitionnistes l'espèce des preuves, qui, certes, est une espèce dérivée, est employée de façon qu'on peut, au cours d'une preuve, opérer avec la supposition de l'existence d'une preuve de quelque assertion – ce qui est une méthode imprédicative. (Bernays, 1962, p. 121) (1.113)

The example of proof being given to show that intuitionism is not wholly predicative, it is clear that Bernays means to flag it as a critical impredicativity. But he does not go on to say that it renders the Proof Explanation non-constructive. On the contrary, upon finding that alternatives (in metamathematics) such as bar induction or computable functionals of finite type likewise introduce impredicativity,[113] he concludes:

> Ainsi nos expériences indiquent que la métamathématique ne peut guère se restreindre dans ses méthodes à des évidences élémentaires ou même seulement prédicatives. (1.114)
>
> Néanmoins nous pouvons maintenir l'idée de la métamathématique et aussi rendre justice à la tendance constructive, cependant nous abstenant dans les méthodes de restrictions innécessaires.
>
> Mais puisqu'il se montre que nous avons à admettre des imprédicativités dans la métamathématique constructive, d'autant moins il y a de raison de rejeter en bloc l'imprédicatif dans les mathématiques classiques. Cela naturellement n'empêche pas que nous tendions généralement à éviter des imprédicativités inutiles.

By the same reasoning, the Proof Explanation would still be considered constructive.

Critical impredicativity of the Proof Explanation is also what Kreisel has in mind when he points out, in his remark (1.4b), that it, like Russell's Paradox, involves a

[112] The debate was opened by Poincaré: 'Ainsi les définitions qui doivent êtres regardées comme non prédicatives sont celles qui contiennent un cercle vicieux. [. . .] Une définition qui contient un cercle vicieux ne définit rien' (Poincaré, 1906, pp. 307, 310). Behmann (1931, pp. 40–41) observed that such definitions are not eliminable, and Beth (1962, p. 83) recalled that the demand that definitions be eliminable had been made in Pascal's 'De l'esprit géométrique et de l'art de persuader' of 1658 (Pascal, 1658/1936). Beth's talk was given at the Colloque international de Mathématiques, Clermont-Ferrand, June 4-7, 1962, organised there at the third centenary of Pascal's death; Bernays' 'Remarques' were presented at the same conference, but he did not rise to the occasion. Gödel (1944/1951, p. 136) pointed out that the circularity poses no problem for realism; on what he thought about its implications for (different kinds of) constructivism, see (1.126) and subsequent discussion below.

[113] For historical context: Bernays writes just after the appearance of Heyting (1956), Kuroda (1956) – see footnote 114 below –, Gödel (1958), Kreisel (1960), and, as he reminiscences to Gödel (Gödel, 2003a, p. 198), had conversations with Spector in Princeton while the latter was working on his (1962); see also the end of footnote 1 of the latter paper.

kind of self-application.[114] The quantification may be construed in two ways. The first is a quantification over all proofs (or over all constructions), as in Kreisel's (1.2). Then $f : A \rightarrow B$ is a proof such that, for all proofs p, if p is a proof of A, then $f(p)$ is a proof of B. The impredicativity is *direct*, in that f is itself among the p. The second is quantification specifically over proofs of A, as in Heyting's clause (1.3). In that case the impredicativity is *indirect*: The informal definition of a proof f of $A \rightarrow B$ then does not quantify over a domain of which f itself is an element, but a proof of A in its domain may contain subproofs of $A \rightarrow B$ and $(A \rightarrow B) \rightarrow A$. The latter subproof is defined in terms of a quantification over proofs of $A \rightarrow B$, among which is f. On either reading, the appearing impredicativity is critical, for lack of a generation procedure for the respective domains containing f.[115] This distinction between direct and indirect impredicativity generalises to other domains and functions.[116]

Dean and Kurokawa (2016, p. 32), quoting Kreisel's (1.4c) but not the preceeding (1.4b), have suggested that a role of (1.4c) precisely is to make one think of the impredicativity of 'the pre-theoretical notion of constructive proof which the BHK interpretation seeks to characterize', by transference from the formal to the informal. And, of course, it will; but I think that the primary reason for following up (1.4b) with (1.4c) was a different one. I will come back to that in my closing remark. In this section, I should like to make some further remarks occasioned by (1.4b) itself.

Kreisel has argued that the impredicativity of the definition of a function f that proves $A \rightarrow B$ may make it impossible to generate the domain of f, but that the constructive acceptability of f does not depend on that possibility to begin with. Rather, we accept f as constructive when inspection of the rule in its definition shows that, whatever we will come to recognise as a constructive proof of A, will, by applying f to it, be turned into a constructive proof of B. A characteristic (and general) passage is:

(1.115) Briefly, to recognize that a given procedure is a welldefined construction, one may already have to have the general notion of construction (similarly in classical mathematics: a formula with quantifiers over sets will in general define a set uniquely only if one already knows the extension of set, except that in the intuitionistic case it is never a matter of the extension). This is an

[114] The first to make this observation about implication (without using the term 'impredicativity') in print seems to have been Kuroda (1956) in his review of Heyting (1956). The first to use that term in print for this seems to have been Kreisel in his review of Wittgenstein (Kreisel, 1958, pp. 147–148). In van Atten (2017b, section 2), I argue that it is not made in Gödel (1933b), as sometimes suggested, but that he did see it as he was working on his functional interpretation.

[115] Incidentally, if there is a proof a of A containing a subproof of $A \rightarrow B$ but not of $(A \rightarrow B) \rightarrow A$, as in skeleton (1.16), then there also is one that contains both, as in skeleton (1.15), provided that the inference steps from the subproof of $A \rightarrow B$ to A remain correct under an open assumption. (Derivation (1.62) in Myhill's Paradox above is one where this is not the case.) The definition of such an a is likewise indirectly impredicative.

[116] An alternative term for indirect impredicativity is 'weak impredicativity', used by Gödel in a note (van Atten, 2015, p. 218) and by Demopoulos (2013, p. 224).

> impredicativity, but constructive, provided, of course, one understands the notion involved. (Kreisel, 1962b, p. 318n8)[117]

Naturally, Kreisel applied this view to intuitionistic species: just as a function gets applied to an object only after that object has been proved to be in its domain, an object becomes an element of a species only by proving that it has the property in question. A species therefore depends as little on a construction method for its elements as a function on a construction method for its domain. The two cases are essentially the same, as to each species corresponds a characteristic function. For the exact definition of species, one would have expected Kreisel to refer to Heyting's *Intuitionism* (1956), which was his reference for the explanations of the logical constants. Heyting's definition runs:

> Definition 1. A species is a property which mathematical entities can be supposed to possess (L. E. J. Brouwer 1918, p. 4; 1924, p. 245; 1952, p. 142). (1.116)
>
> Definition 2. After a species S has been defined, any mathematical entity which has been or might have been defined before S and which satisfies the condition S, is a member of the species S. (Heyting, 1956, p. 37)[118]

But in this case, Kreisel did not follow suit. Without reference to Heyting's definition, in his paper introducing the Theory of Constructions, Kreisel defines

> A *species* of n-tuples of constructions $a_1, \ldots, a_n$ is determined by a construction s where $s(c, a_1, \ldots, a_n) = 0$ if c is a proof that $< a_1, \ldots, a_n >$ belong to the species, $s(c, a_1, \ldots, a_n) = 1$ otherwise. (Kreisel, 1962a, p. 202)[119] (1.117)

and in the Saaty paper he uses the term to name (logically defined) undecided properties: 'called *species* to distinguish them from decidable properties' (Kreisel, 1965, p. 121). One notices immediately the absence of anything corresponding to Heyting's Definition 2 above of a member of a species, which he had elucidated as follows:

> Circular definitions are excluded by the condition that the members of a species S must be definable independently of the definition of S; this condition is obvious from the constructive point of view. It suggests indeed an ordination of species which resembles the hierarchy of types. (Heyting, 1956, p. 38) (1.118)

---

[117] This paper was 'Communicated by Prof. A. Heyting at the meeting of January 27, 1962'. The Heyting-Kreisel correspondence of January 1962 indicates that the final part of that paper was revised quite a bit in the weeks before Heyting presented it to the Academy. Kreisel received the proofs on February 24, and sent the corrected proofs to the Academy on March 8, 1962 (Heyting Papers, Bkre 620308). Gödel's letter to Kreisel quoted in (1.126) below is of the next day.

[118] Below we will have occasion to give, in (1.125), the last of Brouwer's own definitions that Heyting refers to.

[119] The special case for properties of natural numbers had been given in his lecture 'La prédicativité' (Kreisel, 1960, p. 388), held in November 1959, about a year before that of which 'Foundations of intuitionistic logic' (1962) is the published version.

Indeed, Kreisel, in his letter to Heyting of 1962 quoted above in (1.72), had pointed out that 'type distinctions are not always observed, e..g. not in your explanation of the logical constants, in particular implication'. By 1968, he speaks of 'the impredicative theory of species', understanding it as the comprehension principle with intuitionistic logic (Kreisel (1968a, pp. 153); also Kreisel (1968b, p. 351)). Specifically, he proposes to accept as a principle of second-order arithmetic:

$$\exists X \forall y [y \in X \leftrightarrow Ay] , \tag{1.119}$$

where X ranges over species of natural numbers, y ranges over natural numbers, and A may contain quantifiers over species of natural numbers (but not contain the variable X).[120] He elucidates:

(1.120) For Ay to be intuitionistically meaningful, we must have a notion of: *proof of Ay* ([Kreisel 1965], p. 128, 2.31) and this knowledge determines *per se* a species X such that $\forall y(y \in X \leftrightarrow Ay)$.

> What could go wrong? Of course there is the common place objection to impredicative notions allegedly connected with the paradoxes; more precisely we consider here species of arbitrary species instead of sets of arbitrary sets, and take care to derive the paradoxes intuitionistically. Evidently this objection is as weak here as in the case of set theory since we are considering species of natural numbers, and not of arbitrary species. Kreisel (1968a, pp. 153–154)[121]

The paradox arising from accepting species of arbitrary species would be a form of Russell's Paradox for properties (section 1.4.3), but for the species considered here the question of self-membership does not arise. The emphasis on intuitionistic logic in the derivation of the paradox here serves to diagnose that the cause of the paradox lies in the theory of species (see section 1.4.1, and Kreisel's remark (1.72)). Further on, Kreisel comments:

(1.121) Of course it is not claimed that the impredicative species above are our constructions in the sense of our having, so to speak, 'listed' them all before speaking about them, 'listed' in the idealized sense of having given a rule of construction indexed by natural numbers or even ordinals. But note that Heyting's own interpretation of the logical operations, e.g., of implication, certainly does not refer to any 'list' of possible proofs of the antecedent. It simply assumes that we know what a proof is.

> [. . .]

> The moral is not that Heyting's interpretation is non-constructive! nor that a more elementary interpretation such as Gödel's (G) [i.e., the Dialectica interpretation of HA] is foundationally uninteresting. The moral is that its foundational interest depends on something subtler than mere constructive validity. (Kreisel, 1968a, pp. 154–155)

[120] Likewise, Troelstra accepts the existence of a least upper bound on species of the so-called extended reals (Troelstra, 1982, pp. 284–285).

[121] See also Kreisel (1970a, pp. 130-131).

Looking back in 1987 on the period in which (1.115), (1.120), and (1.121) were written, he relates this way of seeing the matter to Gödel and the *Dialectica* paper:

> Asymmetry between rules and – the ranges of – their arguments. One feature that Gödel emphasized increasingly in conversations during the decade after [Gödel 1958] appeared, was the possibility of exploiting the amorphous character – or, if preferred, our ignorance – of the totality of all effective rules. More fully, a rule is accepted only if it is understood to be well defined for all effective arguments (of appropriate type), even though little is – or can be – known about this possibly growing totality. This situation is only superficially paradoxical, to adapt the wording of footnote 1 on p. 283 of [Gödel 1958] about propositional and other logical operators – for the class of propositions – meant by Brouwer and Heyting.[122] (1.122)

The critical impredicativity arising in proof paradoxes when reconstructed with the Fixed Point Theorem is, unlike that in the case of functions, direct, and does not define a species, but an individual object. Three remarks on this point:

1. On the one hand, accepting a representation principle as constructive obliges us to accept its instances as constructive. On the other hand, accepting the critically impredicative definitions in the Proof Paradoxes does not put us in a position actually to construct the defined object. This combination presents us (or rather: we present ourselves) with an obligation that is impossible to fulfill. The infinite regress that Gödel observes in (1.85) is a symptom of this, and parallel regresses arise in each of the other proof paradoxes. The reason can be stated in terms of the proof of the Fixed Point Theorem: the contexts in which these paradoxes arise leave us with no choice for an attempt at constructing the fixed point b but to construct it as the result of bh, to obtain which requires a construction of the fixed point b, and so on.

   This kind of impossibility is not the one we are presented with in the case of mathematical negation, when we observe that our attempt at fulfilling a certain intention directed at a mathematical object with a certain property at some point is blocked, 'no longer goes' (Brouwer 1907, p. 127; trl. Brouwer 1975, p. 73). What we observe here is rather a construction process that never really gets going to begin with. The proposition that such an object exists is therefore non-mathematical.[123]

---

[122] In the footnote referred to, Gödel had written about the concept of computable functional of finite type:

> Man kann darüber im Zweifel sein, ob wir eine genügend deutliche Vorstellung vom Inhalt dieses Begriffs haben, aber nicht darüber, ob die weiter unten angegebenen Axiome für ihn gelten. Derselbe scheinbar paradoxe Sachverhalt besteht auch für den der intuitionistischen Logik zugrunde liegenden Begriff des inhaltlich richtigen Beweises. (Gödel, 1958, p. 283n1) (1.123)

[123] Note that on the Theory of the Creating Subject, a mathematical contradiction can be obtained from this: if for some predicate P the hypothesis that there is a stage n at which $\exists x P(x)$ is proved cannot be true, one concludes that $\neg \exists x P(x)$. The argument uses the implication from

2. The separation of cases between direct and indirect impredicativity shows why the existence of proof paradoxes does not, by analogy, cast doubt on the Proof Explanation of implication. In the proof paradoxes, the direct impredicativity cannot be reconstrued as an indirect one. To do that, it must be possible to see the object defined as a function that is not included in its domain, but in Kreisel's, Gödel's and Goodman's Paradoxes the domain is inclusive of all constructions, and in Troelstra's the object is not a function to begin with. In contrast, in the case of implication both construals are possible: one either quantifies over all constructions, or over the proofs of the antecedent (compare Kreisel's (1.2) and Heyting's (1.3)). This shows that the 'kind of self-application' (Kreisel in (1.4b)) in those paradoxes is not the same as that in implication.[124]
3. Indeed, in the wake of their paradoxes, neither Kreisel, nor Goodman (both of whom sought a mathematical model of the informal Proof Explanation) nor Troelstra (who did not, but *used* it), came to express doubts about the Proof Explanation. Rather, they blocked their paradoxes by (in effect) rejecting the critically impredicative definition coming in with the application, in their particular contexts, of the Fixed Point Theorem: they introduced forms of typing and stratification.

The approach in item 3 is not the one Gödel envisaged to block his intuitionistic paradox. That much is clear from a comment he makes on intuitionism and paradoxes in general, as reported by Wang:

(1.124) Brouwer objects to speaking of all proofs or all constructible objects. Hence the extensional and the intensional paradoxes do not appear in intuitionism according to his interpretation. But I think that this exclusion of all, like the appeal to type theory in the theory of concepts, is arbitrary [from the intuitionistic standpoint]. (Wang, 1996, p. 188, 6.1.15, amendation Wang)

The opening sentence here is, in that exact wording, a half-truth. Brouwer did speak of 'the totality of all possible mathematical systems', as an example of a 'denumerably unfinished set' (Brouwer, 1975, p. 82). A denumerably unfinished set is one of which we can construct a denumerable subset and have a method to extend any such subset; 'from a strictly mathematical point of view this set does not exist as a whole, nor does its power exist; however we can introduce these words here an an expression for a known intention' (Brouwer, 1975, p. 82). (Note that a denumerably unfinished set can be the domain or range of a constructively acceptable morphism.) What Brouwer

right to left in (CS3) on p. 46, or the weaker version $A \to \neg\neg\exists n \Box_n A$, and contraposition, taking $A = \exists x P(x)$. What leads to a block here is that n must be an index into the growing sequence of effected construction acts (which is an object of second-order mathematics), but cannot be. Second-order mathematics is described in Brouwer (1907, p. 98n1) (trl. Brouwer 1975, p. 61n1); see also van Atten (2018, p. 1595n54, p. 1598).

[124] Although, by footnote 3 above, I see no reason to discuss Kreisel's 'second clause' extensively here, I do, coming from a different perspective, express my agreement here with Dean and Kurokawa's view (2016, section 4.2) that Weinstein (1983, p. 264) was mistaken to suggest that it is the self-reflexivity in the 'second clause' that leads to Goodman's Paradox. (See also footnote 100 above.)

objected to is the introduction of a *species* – an existing mathematical object – of all proofs, or of all constructible objects, because he required that the definition of a species take the form of a (not necessarily decidable) separation, and hence be predicative:[125]

> mathematical *species* [are] *properties supposable for mathematical entities previously acquired*, and satisfying the condition that, if they hold for a certain mathematical entity, they also hold for all mathematical entities which have been defined to be equal to it, relations of equality having to be symmetric, reflexive and transitive; mathematical entities previously acquired for which the property holds are called the elements of the species. (Brouwer, 1952, p. 142, original emphasis) (1.125)

The qualification in italics was not present in Brouwer's published definitions before Brouwer (1947),[126] and it can be investigated whether it is a stronger qualification than that imposed on elements in Heyting's 'Definition 2' in (1.116) above.[127] Either way, the ensuing predicativity was also the essence of Brouwer's remarks on denumerably unfinished collections and of his reaction to Russell's Paradox, both formulated in his dissertation (Brouwer, 1907, pp. 148-149, 162–163), which dissertation Gödel had read (van Atten, 2015, p. 191). Once one forbids species with members whose definition is critically impredicative, the proof paradoxes as well as the Russell paradox are all blocked. Surely that was what Gödel had in mind in the first sentence of (1.124), and one begins to wonder whether 'speaking' might be Wang's mishearing or misremembering of Gödel saying 'species'.[128] This would also give a precise sense to Gödel's qualification of arbitrariness, the idea being that if intuitionism accepts impredicativity in the Proof Explanation (as observed, for example, in Bernays' remark (1.113) above) then why not also in the definition of species? Indeed, Brouwer's theory of well-ordered species has generally been considered to be impredicative (see also footnote 138 below).

---

[125] As is evident from the previous footnote, that is not the case for a denumerably unfinished set.

[126] Note that Gödel would have been able to read that, as he read Dutch. But as yet, I have no evidence that he knew the paper.

[127] See on this point van Atten (2017b, section 3).

[128] Wang made notes during and after the conversations, but no tape recordings (Wang, 1996, p. 135). 'Since I do not have a verbatim record of Gödel's own words, there are bound to be misrepresentations' (Wang, 1996, p. 129). The possibility evoked here is reminiscent of Wang's 'Rotterdam' when mentioning Gödel's 1939 lectures at Notre Dame (Wang, 1981, p. 655). (See also Parsons' comment on this in Gödel (2003b, p. 392).)

Gödel insists on the intensional conception of quantification in intuitionism in a letter to Kreisel of March 9, 1962, in a discussion of his Russell paper:

(1.126) Was das vicious circle betrifft, so habe ich ja selbst auf p134 gesagt, dass es sogar für die konstruktive Mathematik nicht in vollen Umfang gilt (vgl. die Formulierung auf p133). Dabei ist ja unter „Konstruktivismus" in meiner ganzen Arbeit der Russell-Poincaré-Weylsche „Halbintuitionismus"[129] zu verstehen, der in einer Hinsicht weiter, in einer andern (insbes⟨esondere⟩ hinsichtlich des Imprädikativen) enger ist als der Intuitionismus. In dem letzteren kommt ja der Begriff der Totalität überhaupt nicht vor u⟨nd⟩ auch die Quantoren sind intensional zu interpretieren (vgl. p136 oben). Es besteht daher kein Grund, weshalb das vic⟨ious⟩ circ⟨le⟩ princ⟨iple⟩ (intensional formuliert) im Intuit⟨ionismus⟩ gelten sollte. (Kreisel Papers, 50/1, underlining Gödel)[130]

In the paper, Gödel had stated the vicious circle principle as 'no totality can contain members definable only in terms of this totality, or members involving or presupposing this totality', and remarked that, for each choice among 'definable only in terms of', 'involving', and 'presupposing', one in fact obtains a different principle (Gödel, 1944/1951, pp. 133, 135). The intensional formulation is that in terms of definability. In it, intuitionism would replace the extensional 'totality' by the intensional species, but then Gödel sees no objection to a species having members that are definable only in terms of that species. Kreisel's acceptance of impredicative definitions as in (1.115)–(1.122) above is based on this idea.

Not long after that letter, Gödel made the point in print, in a note he added to the 1964 reprint of that paper:[131]

(1.127) The author wishes to note [...] that the term 'constructivistic' in this paper is used for a strictly anti-realistic kind of constructivism. Its meaning, therefore, is not identical with that used in current discussions on the foundations of mathematics. If applied to the actual development of logic and mathematics it is equivalent with a certain kind of 'predicativity' and hence different both from 'intuitionistically admissible' and from 'constructive' in the sense of the Hilbert School. (Benacerraf & Putnam, 1964, p. 211)

[129] Die finite Mathematik scheint mir in einem gewissen Sinn der Durchschnitt von Intuition⟨ismus⟩ u⟨nd⟩ Halbintuit⟨ionismus⟩ zu sein.

[130] In the letter of January 31, 1962 to which Gödel is replying, Kreisel (letter of January 31, 1962 (Kreisel Papers, 50/2)) had said that he had read the Russell paper again, and found the argument there that the vicious circle principle should apply to constructive mathematics (Gödel, 1944/1951, pp. 136–137) not convincing, on the ground of the considerations in the *Dialectica* paper. But that is, as we see, not the kind of constructivism Gödel had meant.

[131] He expanded and revised that note in 1972, and only that version is included in Gödel (1990). But see the next footnote for a related note that is included there.

Gödel here sees, by implication, an element (not a 'vestige'!) of realism in the intuitionistic position;[132] and we see that (1.126) and (1.127) confirm the correctness of Wang's amendation to (1.124).

The paradoxes that drew attention to the problematic character of the impredicative species of all proofs or all constructions – the proof paradoxes of section 1.4.4 – all appeared in print, and were discussed among the protagonists, over a period of a few years after (1.127). In 1969, Gödel stated where he wanted to look for a solution to paradoxes, in a letter to Kreisel (July 25):

> Die Scott-schen Bedenken gegen impräd⟨ikativen⟩ Spezies u⟨nd⟩ den allgemeinen Beweisbarkeitsbegriff[133] scheinen mir *beim heutigen Stand der Wissenschaft* durchaus berechtigt.[134] Wie die Antinomien zeigen,[135] *verstehen* wir diese sehr allgemeinen Begriffe heute noch *nicht*. Erst nach einer genauen phänomenolog⟨ischen⟩ Analyse, welche die Antinomien auf eine vollkommen einleuchtende Weise auflöst, werden sie vertrauenerweckend sein. (Kreisel Papers, 50/2, emphasis Gödel) (1.129)

These lines were written at the time of Gödel's undergirding his Dialectica Interpretation by Husserl's phenomenology, first, more specifically, as a contribution to intuitionism, then in the form of a theory of 'reductive proof' (van Atten, 2015, pp. 210–222). This is no coincidence, of course: philosophical questions around impredicativity and proof are raised just as much by the concept of computable functional; see, e.g., Gödel's (1.123) in footnote 122 above. One thinks of Husserl: 'Für die apriorischen Disziplinen, die innerhalb der Phänomenologie zur Begründung kommen (z.B. als mathematische Wissenschaften) [kann es] keine ‚Paradoxien', keine ‚Grundlagenkrisen' geben' (Husserl, 1962, p. 297).[136]

---

[132] In a note kept with an offprint of the Russell paper, Gödel specified an 'antireal⟨istic⟩ kind of constr⟨uctivism⟩' as one for which

> the starting point and the means of constr⟨uction⟩ are to be exclusively sensual & material (e.g. symbols, their perc⟨eptual⟩ prop⟨erties⟩ & rel⟨ations⟩ and the actual or imagined handling of them), not the element⟨ary⟩ operations and int⟨uitions⟩ of a new & irreducible entity called mind. (Gödel, 1990, p. 320) (1.128)

In Sundholm and van Atten (2008, p. 71), intuitionism and platonism were likened to one another as forms of 'ontological descriptivism' (there, as opposed to meaning-theoretical approaches).

[133] This refers to Scott's thinking, which he discussed with Gödel and Kreisel, that led to „Constructive validity" (1970); see its p. 239 and p. 241 (where, however, no explicit argument against quantification over species is given). See also footnote 140 below.

[134] [Note MvA] Kreisel comments: 'Scott, in [Scott 1970], p. 239, 1.-10 to 1.-9, expresses very clearly similar misgivings about the role of proofs in constructive foundations. Pushed beyond reason [...] Scott's view blocks any chance, at least at present, of a non-circular explanation of implication.' (Kreisel, 1971c, p. 124n8)

[135] [Note MvA] Here Gödel may also have had paradoxes such as Myhill-Montague in mind (see the end of section 1.4.1).

[136] On his 'reductive proof', which Gödel introduced around the time of (1.129), he observed in a note dated February 11, 1974:

> <u>Meine Dial⟨ectica⟩ Arbeit mit dem Begriff des reduktiven Beweis⟨es⟩ gibt keine die Parad⟨oxien⟩</u> ausschließende Interpretation (daher <u>die Fundierung</u> nicht wesentlich bes- (1.130)

Gödel's expectation in (1.129) is that an analysis such as he envisages will have the effect of validating impredicative species and a general notion of proof.[137] Brouwer, on the other hand, at least in his explicit statements, and presumably on the basis of *his* further analysis, accepts a general concept of proof, but not impredicatively defined species, e.g. (1.125) above. It has been a matter of debate (*i*) whether in Brouwerian intuitionism, as practised, critical impredicativities nevertheless do occur;[138] (*ii*) whether, given Brouwer's views on mathematical existence and truth, they should be avoided; and (*iii*) whether, if they should, intuitionism would still be able to develop satisfactory understandings of, in particular, implication and of well-orderings. A treatment of those questions lies outside the scope of the present paper; for references, a discussion, and an attempt at a contribution, see van Atten (2017b).

As for Kreisel, already at the time of receiving Gödel's letter with (1.129) in it, his thinking was developing in quite the opposite (deflationary) direction, in every respect. Having given his analysis of impredicative species quoted in (1.90) above, he recommended leaving it at that:

(1.131) The analysis above, like the interpretation of the logical operations intended by Brouwer and formulated by Heyting, uses notions which are more abstract than those of familiar constructive mathematics [. . .] The analysis has enough coherence and substance to suggest that there is something deflnite to understand here [. . .] But do we want to know about it, not only subjectively, but for getting on with the business of constructive mathematics? Not the possibility of understanding intuitionistic concepts, but their usefulness is the true issue. Dramatic exaggerations would only lead to the kind of let-

---

ser als Heyting und zwar deswegen, weil zum Beispiel der allgemeine Begriff der berechenbaren zahlentheoretischen Funktion vorkommt und dieser von irgendeiner Def⟨initions⟩-Kette spricht (also die Def⟨inition⟩ $x \in a \equiv {\sim}x \in x$ kann vorkommen). (Gödel Papers, 10a/40, item 050136, transcription from the Gabelsberger Eva-Maria Engelen.))

This note was first published, and discussed, in van Atten (2015, p. 222–224).

[137] A view that I cannot go into here, but that is highly interesting for comparison and contrast with Gödel's, is that of those who wished to accept both impredicative species and the Curry-Howard isomorphism. This combination was proved inconsistent by Girard (1972); see also Coquand (1986). Generally, the conclusion was that impredicative species had to go (Martin-Löf, 2008, p. 250), and this opened the way to Martin-Löf's Constructive Type Theory as we know it today. Gödel, on the contrary, had accepted, from early on, as the object correlate to a general notion of proof, one universe of all proofs. He (and Kreisel) criticised Howard's manuscript for not analysing what a construction is (Wadler, 2014, p. 11). (Incidentally, Bill Howard told me that Gödel and he never discussed impredicativity in the Proof Explanation explicitly (van Atten, 2015, p. 193n10).)

[138] In his notebook Max Phil IV of May 1941-April 1942, Gödel lists as impredicativities in Brouwer the sum species of an arbitrary species (of species) and the definition of ordinal numbers (Gödel Papers, 6b/67, item 030090, p. 154). See, e.g., Brouwer's definition of the 'Vereinigungsspecies' (and note the use of $\mathfrak{S}$, as for 'Summe') in Brouwer (1925, p. 247), and the definition of well-ordered species, which is fundamental to Brouwer's theory of ordinal numbers, in Brouwer (1927b, pp. 451, 456). In the former definition, the implicit quantification over species is not universal but existential. Note that Gödel does not mention implication; perhaps because in Brouwer the Proof Explanation is operative, but implicit (van Atten, 2017a, section 3.1).

> down which Russell felt after he (or rather, according to his autobiography, after Whitehead) finished Principia. (Kreisel, 1970a, p. 131)[139]

In fact, Gödel's (1.129) was written *after* he had seen this recommendation, and no doubt (also) records his reaction to it.[140] And while Kreisel initially had come to see, no doubt because of their conversations, a role for phenomenological analysis in foundations just as Gödel does in (1.129) – see Kreisel 1969, p. 97; 1970b, p. 489; and 1971b, p. 151 – such calls are absent from his later writings.[141]

Finally, as we saw in (1.81), a comment of 1958 and among Kreisel's first on the paradoxes, he looked at Russell's and concluded that it arises from an oversight: not every property has a definite extension. He there also stated that this take, while 'illuminating', 'does not seem to lend itself to generalisation'. It seems that initially he nevertheless hoped that such a generalisation would be found; thus in 1967 he writes that

[139] The qualification 'dramatic exaggeration' predicts Kreisel's reaction to the elaboration of the *Dialectica* paper that Gödel was working on at that very moment. Kreisel always considered the 1958 version a 'gem' (Kreisel, 1987b, p. 108), but, reviewing volume 2 of the *Collected Works*, considers the added notes in the 1972 version 'particularly ethereal', and opines that, while it is not difficult to see the philosophical gain achieved 'when the "gain" is measured by the canons of academic epistemology', '[t]his leaves open what gain, if any, there is for a more realistic view (of knowledge)' (Kreisel, 1990, p. 615).

[140] The details are these. From February 20–25, 1969, Scott had written (and then sent) Gödel a long letter further clarifying his (Scott's) theory of constructions . This included reservations about quantification over species, e.g.,

> A distinction is possible in intuitionistic logic that is *not* possible in classical logic – namely, functions are *sharp* while species are *fuzzy*. (At the moment I cannot find less colloquial language to express this point.) That makes it possible to say that we *can* imagine the *whole* function space $\mathbb{N} \to \mathbb{N}$ but not the whole space of subspecies of N. (Of course, $\mathbb{N} \to 2$ gives the space of *detachable* subspecies of N, but we are also interested in the *undecidable* subspecies.) Somehow the introduction of new mathematical objects (higher-type objects, say) allows for the formulation of ever new properties of integers making it *impossible* to comprehend all subspecies into *one* totality. That is a vague idea, I admit, but I would like to see how you or Kreisel can argue that it is unreasonable. (Gödel Papers, 3a/155, item 012275, pp. 28–29) (1.132)

(I thank Dana Scott for his permission to quote.) Scott sent a copy to Kreisel. I have not seen a responding letter from Kreisel to Scott, but a letter from Kreisel to Gödel of March 10, 1969 (Gödel Papers, 2a/92, item 011262) shows that Kreisel then asked Myhill, one of the editors of the Buffalo volume, to send Gödel the typescript of Kreisel (1970a). In that letter, Kreisel commented that at least some of Scott's objections seem to be incorrect, for the reasons given in section 5 of that paper (without there naming Scott). 'Übrigens sind Ihnen die Überlegungen von §5 höchstwahrscheinlich sowieso nicht fremd' (a reference, I take it, to the conversations Kreisel reports in (1.122)). Gödel gave his view on Scott's reservations to Kreisel first – this is (1.129) above – and only on December 18 to Scott, in a similar but somewhat more practical manner: 'I share your strongly intensional viewpoint as to intuitionistic mathematics and (pending further clarification) your distrust in unlimited functions, and quantification over all (intensional) subspecies, or over all proofs or propositions. I think it makes very good sense first to develop a theory of functions (i.e., procedures) with limited domains, but including transfinite types.' (Gödel Papers, 3a/155, item 012279, p. 1).

[141] As he explained later, he did not read Husserl because he was, in fact, 'interested in other things' (1998, pp. 100, 105).

(1.133) in Zermelo's work [. . .] the intuitive analysis of the crude mixture of notions, namely the description of the type structure, led to the good axioms [. . .] And a similar conceptual analysis will be needed for solving the problem of the paradoxes. (Kreisel, 1967a, p. 145)

But by 1971, at the beginning of a long '*Autobiographical remark* on the (functional) paradoxes', a turn sets in:

(1.134) Speaking for myself, I simply do not find the paradoxes dramatic: halfway through the argument, that is well before any hint of a paradox appears, my attention begins to wander as in free association. (Kreisel, 1971b, p. 188)

And in 1973 he holds:

(1.135) For contrary to popular opinion I have the impression that paradoxes occur when we have not even *begun* to think, when we are playing with words, and their resolution is generally *not* fruitful: after all, how much more does the child really know about the concept of number when he realizes that there is no greatest? (Kreisel, 1973, p. 265)

Not fruitful, unlike, one is, in view of the above, inclined to add, Zermelo's analysis of set.[142] In his retrospective Salzburg essay, he makes the contrast with Gödel explicit:

(1.138) Evidently, if such simple and familiar points are overlooked in the manufacture of paradoxes, there is good reason to doubt Gödel's high expectations from a solution of the paradoxes. (Kreisel, 1987a, pp. 95-96)

The change in Kreisel's view on the paradoxes described here happened soon after he had written the Brouwer memoir, and I take this to explain why he seems never to have taken up the matter of (1.4b) again.

---

[142] Kreisel (1967a, p. 144n1) still expressed his opposition to the view, there ascribed to Rasiowa and Sikorski (1963), of the paradoxes as a 'dead (fruitless) issue'. Also note how close the spirit of Kreisel's (1.135) to that of this passage in Brouwer:

(1.136) It can be shown, however, that these paradoxes result from the same error as that of Epimenides, namely, that they arise where regularities in the language that accompanies mathematics are extended over a language of mathematical words that does not accompany mathematics; that, further, logistics too is concerned with the mathematical language instead of with mathematics itself, thus does not clarify mathematics itself; that, finally, all paradoxes disappear, when one restricts oneself to speaking only of systems that explicitly can be built out of the Ur-intuition. [Brouwer 1908, p. 155; trl. van Atten and Sundholm 2017, p. 17]

Weyl, too, has drawn this contrast, in his remarks on Gödel (1944/1951):

(1.137) [O]ne must know that Gödel takes the paradoxes very seriously; they reveal to him 'the amazing fact that our logical intuitions are self-contradictory'. This attitude towards the paradoxes is of course at complete variance with the view of Brouwer who blames the paradoxes not on some transcendental logical intuition which deceives us, but on a gross error inadvertently committed in the passage from finite to infinite sets. I confess that in this respect I remain steadfastly on the side of Brouwer. (Weyl, 1946, p. 211)

1.5 Closing remarks

The question why Kreisel, to the best of my knowledge, never specified what A he had in mind in (1.4c), is one that only he could have answered. What I have tried here is to find examples that Kreisel knew in any case for other reasons, or that occur in contexts that he was particularly familiar with.

A question I need to return to is that of the relation between remarks (1.4b) and (1.4c). As mentioned at the beginning of section 1.4.5, I agree with Dean and Kurokawa that (1.4c) will draw attention to the idea of impredicativity of implication, which is a way in which proofs are about themselves, as stated in (1.4b). However, Kreisel could have done that in a more direct way, without also invoking the contrast between the formal and the non-formal, or that between the natural and the non-natural. The point of inserting the parenthetical remark (1.4c), such as I understand it, will also serve as conclusion to the present paper as a whole. It is to make clear that for certain A, a formal proof that depends on a proof of $A \to B$ is not a merely theoretical, and in that sense unnatural possibility, to be taken into account only to test the generality of some theoretical interpretation or explanation of the pre-theoretical notion 'constructive proof of A', but one that, when formalising natural informal proofs, is to be expected.

Acknowledgements. I am grateful to Göran Sundholm for his teaching, collaboration, conversation, and friendship over the years.

For exchanges on the topics or sources while writing this paper, I thank Henk Barendregt, Gabriella Crocco, Walter Dean, Ken Derus, Dan Isaacson, Ansten Klev, Hidenori Kurokawa, Gianluca Longa, Reviel Netz, Jan von Plato, Adrian Rezuş, Andrei Rodin, Robin Rollinger, Göran Sundholm, Dana Scott, Wim Veldman, and Albert Visser. I thank the reviewer for comments and questions on the first version I handed in.

Gödel's unpublished writings are copyright by the Institute for Advanced Study, Princeton, and located in (*i*) the Kurt Gödel Papers at the Shelby White and Leon Levy Archives center at the IAS, on deposit at the Manuscripts Division, Department of Special Collections, Princeton University Library; and (*ii*) the Georg Kreisel Papers at the Department of Special Collections of the Stanford University Archives.

Kreisel's unpublished writings are copyright by the Estate of Georg Kreisel (represented by Professor Paul Weingartner, Salzburg), and on deposit at the Department of Special Collections of the Stanford University Archives, and, as part of the Arend Heyting Papers, at the Noord-Hollands Archief, Haarlem.

The quotations in this paper appear by permission of the rights holders, to whom I make grateful acknowledgement.

For assistance with the sources in these archives, I thank Leif Anderson and Tim Noakes (both at Stanford), Hans van Felius (Haarlem), Marcia Tucker (Institute for Advanced Study, Princeton), Brianna Cregle, Charles Doran, and AnnaLee Pauls (all three at the Princeton University Library), Paul Weingartner (Salzburg), and also

Daniel Wilhelm of the Philosophische Archiv at the Universität Konstanz, which holds another part of Kreisel's archive.

References

Ackermann, W. (n.d.). Review of Behmann (1959). *Zentralblatt für Mathematik und ihre Grenzgebiete*, *Zbl 0095.24201*.

Artemov, S. (2001). Explicit provability and constructive semantics. *Bulletin of Symbolic Logic*, *7*(1), 1–36.

Barendregt, H. (1984). *The lambda calculus: Its syntax and semantics* (2nd, revised ed.). North-Holland.

Behboud, A. (1994). Greek geometrical analysis. *Centaurus*, *37*(1), 52–86.

Behmann, H. (1931). Zu den Widersprüchen der Logik und der Mengenlehre. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, *40*, 37–48.

Behmann, H. (1959). Der Prädikatenkalkül mit limitierten Variabelen: Grundlegung einer natürlichen exakten Logik. *Journal of Symbolic Logic*, *24*(2), 112–140.

Benacerraf, P., & Putnam, H. (Eds.). (1964). *Philosophy of Mathematics: Selected Readings*. Cambridge University Press.

Bernays, P. (1947). Review of Pólya (2004) [1945]. *Dialectica*, *1*(2), 178–188.

Bernays, P. (1962). Remarques sur l'imprédicativité. *Annales scientifiques de l'Université de Clermont-Ferrand 2*, *tome 7, série Mathématiques*(1), 121–122.

Beth, E. W. (1962). Observations concernant la théorie de la définition. *Annales scientifiques de l'Université de Clermont-Ferrand 2*, *tome 7, série Mathématiques*(1), 83–87.

Boolos, G. (1984). Don't eliminate cut. *Journal of Philosophical Logic*, *13*(4), 373–378.

Brouwer, L. E. J. (1907). *Over de grondslagen der wiskunde* (Doctoral dissertation). Universiteit van Amsterdam.

Brouwer, L. E. J. (1908). De onbetrouwbaarheid der logische principes. *Tijdschrift voor Wijsbegeerte*, *2*, 152–158.

Brouwer, L. E. J. (1925). Zur Begründung der intuitionistischen Mathematik I. *Mathematische Annalen*, *93*, 244–257.

Brouwer, L. E. J. (1927a). Über Definitionsbereiche von Funktionen. *Mathematische Annalen*, *97*, 60–75.

Brouwer, L. E. J. (1927b). Zur Begründung der intuitionistischen Mathematik III. *Mathematische Annalen*, *96*, 451–488.

Brouwer, L. E. J. (1947). Richtlijnen der intuïtionistische wiskunde. *Indagationes Mathematicae*, *9*, 197.

Brouwer, L. E. J. (1952). Historical background, principles and methods of intuitionism. *South African Journal of Science*, *49*, 139–146.

Brouwer, L. E. J. (1975). *Collected works*. *Vol. 1*. *Philosophy and foundations of mathematics* (A. Heyting, Ed.; *Vol. 1*). North-Holland.

Buldt, B. (1997). Review of Kurt Gödel, On formally undecidable propositions of Principia Mathematica and related systems, Translated by B. Meltzer, with an introduction by R. B. Braithwaite. *Review of Modern Logic*, *7*(1), 90–99.

Church, A. (1932). A set of postulates for the foundation of logic [I]. *Annals of Mathematics. Second Series*, *33*, 346–366.

Church, A. (1933). A set of postulates for the foundation of logic [II]. *Annals of Mathematics. Second Series*, *34*, 839–864.

Church, A. (1935–1936). Mathematical Logic [Unpublished lecture notes. The consulted copy is a scan of the one in Curry's private library, obtained by A. Rezuş from J. Seldin. Another copy is availaible in the Church Papers, C0948, Dpeartment of Special Collections, Princeton University Library, box 10, folder 2].

Church, A. (1941). *The calculi of lambda-conversion* [Reprint 1951]. Princeton University Press, Princeton, N. J.

Coquand, T. (1986). An analysis of Girard's Paradox. *LICS 1986*, 227–236.

Crocco, G. (2019). Informal and absolute proofs: Some remarks from a Gödelian perspective. *Topoi*, *38*, 561–575.

Curry, H. B. (1934). Functionality in combinatory logic. *Proceedings of the National Academy of Sciences of the United States of America*, *20*, 584–590.

Curry, H. B. (1941). Some aspects of the problem of informal rigor. *Bulletin of the American Mathematical Society*, *47*, 221–241.

Curry, H. B. (1942). The inconsistency of certain formal logics. *Journal of Symbolic Logic*, *7*, 115–117.

D'Agostino, M., Finger, M., & Gabbay, D. (2008). Cut-based abduction. *Logic Journal of the IGPL*, *16*(6), 537–560.

D'Agostino, M., & Mondadori, M. (1994). The taming of the cut: Classical refutations with analytic cut. *Journal of Logic and Computation*, *4*(3), 285–319.

Davis, M. (Ed.). (1965). *The undecidable*. Raven Press.

Dawson, C. (1984). *Books and journals in the Nachlass of Kurt Gödel* [Unpublished].

Dean, W. (2014). Montague's paradox, informal provability, and explicit modal logic. *Notre Dame Journal of Formal Logic*, *55*(2), 157–196.

Dean, W. (2020). Incompleteness via paradox and completeness. *Review of Symbolic Logic*, *13*(3), 541–592.

Dean, W., & Kurokawa, H. (2016). Kreisel's Theory of Constructions, the Kreisel-Goodman Paradox, and the Second Clause. In T. Piecha & P. Schroeder-Heister (Eds.), *Advances in proof-theoretic semantics* (pp. 27–63). Springer.

Demopoulos, W. (2013). *Logicism and its philosophical legacy*. Cambridge University Press.

Derus, K. (2020). Chitchat with the Devil: Kreisel's letters, 2002–2015. In P. Weingartner & H.-P. Leeb (Eds.), *Kreisel's interests* (pp. 121–171). College Publications.

Diaconescu, R. (1975). Axiom of choice and complementation. *Proceedings of the American Mathematical Society*, *51*, 176–178.

Dijkstra, E. (1985, September). On anthropomorphism in science (EWD936). http://www.cs.utexas.edu/users/EWD/ewd09xx/EWD936.PDF

Escardó, M., & Oliva, P. (2010). Selection functions, bar recursion and backward induction. *Mathematical Structures in Computer Science*, *20*(2), 127–168.

Euclid. (1908). *The Thirteen Books of* The Elements. *Vol. 1. Books I and II* (T. Heath, Ed.; *Vol. 1*). Cambridge University Press.

Fenstad, J. (Ed.). (1971). *Proceedings of the second Scandinavian logic symposium*. North-Holland.

Fitch, F. (1952). *Symbolic logic: An introduction*. The Ronald Press Co.

Fitting, M. (2008). A quantified logic of evidence. *Annals of Pure and Applied Logic*, *152*(1–3), 67–83.

Freudenthal, H. (1936). Zur intuitionistischen Deutung logischer Formeln. *Compositio Mathematica*, *4*, 112–116.

Geach, P. (1955). On insolubilia. *Analysis*, *15*(3), 71–72.

Girard, J.-Y. (1972). *Interprétation fonctionelle et élimination des coupures dans l'arithmétique d'ordre supérieure* (Doctoral dissertation). Université Paris 7.

Gödel, K. (1932). Review of Church (1932). *Zentralblatt für Mathematik und ihre Grenzgebiete*, *Zbl 0004.14507*.

Gödel, K. (1933a). Eine Interpretation des intuitionistischen Aussagenkalküls. *Ergebnisse eines mathematischen Kolloquiums*, *4*, 39–40.

Gödel, K. (1933b). The present situation in the foundations of mathematics [First published in Gödel (1995, pp. 45–53)].

Gödel, K. (1934). Review of Church (1933). *Zentralblatt für Mathematik und ihre Grenzgebiete*, *Zbl 0008.28902*.

Gödel, K. (1946). Remarks before the Princeton bicentennial conference on problems in mathematics [First published in Davis (1965), pp. 84–88].

Gödel, K. (1951). Russell's mathematical logic. In P. Schilpp (Ed.), *The Philosophy of Bertrand Russell* (3rd ed., pp. 123–153). Tudor. (Original work published 1944)
Gödel, K. (1958). Über eine bisher noch nicht benutzte Erweiterung des finiten Standpunktes. *Dialectica*, *12*, 280–287.
Gödel, K. (1972). On an extension of finitary mathematics which has not yet been used. In S. Feferman, J. Dawson Jr, S. Kleene, G. Moore, R. Solovay, & J. van Heijenoort (Eds.), *Collected works*. *Vol. 2*. *Publications 1938–1974*. Oxford University Press.
Gödel, K. (1990). *Collected works*. *Vol. 2*. *Publications 1938–1974* (S. Feferman, J. Dawson Jr, S. Kleene, G. Moore, R. Solovay, & J. van Heijenoort, Eds.; *Vol. 2*). Oxford University Press.
Gödel, K. (1995). *Collected works*. *Vol. 3*. *Unpublished essays and lectures* (S. Feferman, J. Dawson Jr, W. Goldfarb, C. Parsons, & R. Solovay, Eds.; *Vol. 3*). Oxford University Press.
Gödel, K. (2003a). *Collected works*. *Vol. 4*. *Correspondence A-G* (S. Feferman, J. Dawson Jr, W. Goldfarb, C. Parsons, & W. Sieg, Eds.; *Vol. 4*). Oxford University Press.
Gödel, K. (2003b). *Collected works*. *Vol. 5*. *Correspondence H-Z* (S. Feferman, J. Dawson Jr, W. Goldfarb, C. Parsons, & W. Sieg, Eds.; *Vol. 5*). Oxford University Press.
Gödel Papers. C0282, Department of Special Collections, Princeton University Library. Cited by box/folder and item number.
Goodman, N. (1970). A theory of constructions equivalent to arithmetic: Proceedings of the summer conference at Buffalo N. Y., 1968 (A. Kino, J. Myhill, & R. Vesley, Eds.), 101–120.
Goodman, N. (1973). The arithmetic theory of constructions (H. Rogers & A. Mathias, Eds.). (337), 274–298.
Goodman, N. (1979). Mathematics as an objective science. *American Mathematical Monthly*, *86*, 540–551.
Goodman, N., & Myhill, J. (1978). Choice implies excluded middle. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, *24*(5), 461.
Hacking, I. (2014). *Why Is There Philosophy of Mathematics a All?* Cambridge University Press.
Hankel, H. (1874). *Zur Geschichte der Mathematik in Alterthum und Mittelalter*. B. G. Teubner.
Heyting, A. (n.d.). Review of Goodman (1970). *Mathematical Reviews*, *MR 282797*.
Heyting, A. (1956). *Intuitionism: An introduction*. North–Holland.
Heyting Papers. Noord-Hollands Archief, Haarlem. Cited by item number.
Hintikka, J., & Remes, U. (1976). Ancient geometrical analysis and modern logic. In R. Cohen (Ed.), *Essays in memory of Imre Lakatos* (pp. 253–276). Reidel.
Husserl, E. (1962). *Husserliana*. *Vol. IX*. *Phänomenologische Psychologie* (W. Biemel, Ed.; *Vol. IX*). Martinus Nijhoff.
Isaacson, D. (2020). Georg Kreisel: Some biographical facts. In P. Weingartner & H.-P. Leeb (Eds.), *Kreisel's interests* (pp. 87–119). College Publications.
Jeroslow, R. (1973). Redundancies in the Hilbert-Bernays derivability conditions for Gödel's second incompleteness theorem. *Journal of Symbolic Logic*, *38*, 359–367.
Kleene, S. (1934). Proof by cases in formal logic. *Annals of Mathematics. Second Series*, *35*, 529–544.
Kleene, S. (1945). On the interpretation of intuitionistic number theory. *Journal of Symbolic Logic*, *50*, 109–124.
Kleene, S. (1952). *Introduction to metamathematics*. North-Holland.
Kleene, S. (1981). Origins of recursive function theory. *Annals of the History of Computing*, *3*(1), 52–67.
Knorr, W. (1983). Construction as existence proof in ancient geometry. *Ancient Philosophy*, *3*, 125–148.
Knorr, W. (1986). *The ancient tradition of geometric problems*. Birkhäuser.
Kolmogorov, A. (1932). Zur Deutung der intuitionistischen Logik. *Mathematische Zeitschrift*, *35*, 58–65.
Kolmogorov, A. (2019). *Selected works*. *Vol. 1*. *Mathematics and mechanics* (V. Tikhomirov, Ed.; V. Volosov, Trans.; *Vol. 1*). Springer.

Kosheleva, O., & Kreinovich, V. (2015). Constructive mathematics is seemingly simple but there are still open problems: Kreisel's observation explained. *Journal of Innovative Technology and Education*, *2*(1), 51–56.
Kreisel, G. (n.d.-a). Review of Goodman (1970). *Zentralblatt für Mathematik und ihre Grenzgebiete*, *Zbl 0212.31301*.
Kreisel, G. (n.d.-b). Review of Goodman (1973). *Zentralblatt für Mathematik und ihre Grenzgebiete*, *Zbl 0272.02040*.
Kreisel, G. (n.d.-c). Review of Löb (1955). *Mathematical Reviews*, *MR 70596*.
Kreisel, G. (n.d.-d). Review of Smullyan (1965). *Mathematical Reviews*, *MR 200135*.
Kreisel, G. (n.d.-e). Review of Weinstein (1983). *Zentralblatt für Mathematik und ihre Grenzgebiete*, *Zbl 0539.03037*.
Kreisel, G. (1958). Wittgenstein's *Remarks on the Foundations of Mathematics*. *British Journal for the Philosophy of Science*, *9*, 135–158.
Kreisel, G. (1960). La prédicativité. *Bulletin de la Société mathématique de France*, *88*, 371–391.
Kreisel, G. (1962a). Foundations of intuitionistic logic. In E. Nagel, P. Suppes, & A. Tarski (Eds.), *Logic, methodology, and philosophy of science, proceedings of the international congress, stanford 1960* (pp. 198–210).
Kreisel, G. (1962b). The axiom of choice and the class of hyperarithmetic functions. *Koninklijke Nederlandse Akademie van Wetenschappen, Proceedings, Series A*, *65*, 307–319.
Kreisel, G. (1965). Mathematical logic. In T. Saaty (Ed.), *Lectures on modern mathematics* (pp. 95–195). Wiley.
Kreisel, G. (1967a). Informal rigour and completeness proofs. In I. Lakatos (Ed.), *Problems in the Philosophy of Mathematics* (138–186). North-Holland.
Kreisel, G. (1967b). Mathematical logic: What has it done for the philosophy of mathematics? Philosopher of the century. In R. Schoenman (Ed.), *Bertrand Russell* (201–272). George Allen; Unwin.
Kreisel, G. (1968a). Functions, ordinals, species. In B. van Rootselaar & J. Staal (Eds.), *Logic, methodology, and philosophy of science, proceedings of the 3rd international congress, amsterdam 1967* (pp. 145–159).
Kreisel, G. (1968b). A survey of proof theory. *Journal of Symbolic Logic*, *33*, 321–388.
Kreisel, G. (1969). Two notes on the foundations of set-theory. *Dialectica*, *23*, 93–114.
Kreisel, G. (1970a). Church's thesis: A kind of reducibility axiom for constructive mathematics. In A. Kino, J. Myhill, & R. Vesley (Eds.), *Intuitionism and proof theory: Proceedings of the summer conference at Buffalo N. Y., 1968* (pp. 121–150). North-Holland.
Kreisel, G. (1970b). Principles of proof and ordinals implicit in given concepts. In A. Kino, J. Myhill, & R. Vesley (Eds.), *Intuitionism and proof theory: Proceedings of the summer conference at Buffalo N. Y., 1968* (pp. 489–516). North-Holland.
Kreisel, G. (1970c). The formalist-positivist doctrine of mathematical precision in the light of experience. *L'Âge de la Science*, *3*, 17–46.
Kreisel, G. (1971a). Review of *The Collected Papers of Gerhard Gentzen*. Ed. M. E. Szabo. *Journal of Philosophy*, *68*(8), 238–265.
Kreisel, G. (1971b). Some reasons for generalizing recursion theory. In *Logic Colloquium '69: Proceedings of the summer school and colloquium, Manchester, 1969* (pp. 139–198).
Kreisel, G. (1971c). A survey of proof theory [II]. In J. Fenstad (Ed.), *Proceedings of the second Scandinavian logic symposium* (pp. 109–170). North-Holland.
Kreisel, G. (1972). Which number theoretic problems can be solved in recursive progressions on Π^1_1-paths through O? *Journal of Symbolic Logic*, *37*(2), 311–334.
Kreisel, G. (1973). Perspectives in the philosophy of pure mathematics. *Logic, Methodology, and Philosophy of Science, Proceedings of the 4th International Congress, Bucharest 1971*, 255–277.
Kreisel, G. (1976). *Postscript to the review of Gentzen's Collected Papers* [Unpublished. Typescript copy by Göran Sundholm].
Kreisel, G. (1983). Hilbert's programme. In P. Benacerraf & H. Putnam (Eds.), *Philosophy of mathematics: Selected readings* (2nd ed., pp. 207–238). Cambridge University Press.

Kreisel, G. (1987a). Church's thesis and the ideal of informal rigour. *Notre Dame Journal of Formal Logic*, *28*(4), 499–519.
Kreisel, G. (1987b). Gödel's excursions into intuitionistic logic. In P. Weingartner & L. Schmetterer (Eds.), *Gödel remembered* (67–179). Bibliopolis.
Kreisel, G. (1989). Zu einigen Gesprächen mit Wittgenstein: Erinnerungen und Gedanken. In *Wittgenstein: Biographie-Philosophie-Praxis* (pp. 131–143). Wiener Secession.
Kreisel, G. (1990). Review of Gödel's 'Collected Works, Volume II'. *Notre Dame Journal of Formal Logic*, *31*(4), 602–641.
Kreisel, G. (1998). Second thoughts around some of Gödel's writings: A non-academic option. *Synthese*, *114*(1), 99–160.
Kreisel, G., & Lévy, A. (1968). Reflection principles and their use for establishing the complexity of axiomatic systems. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, *14*, 97–142.
Kreisel, G., & Macintyre, A. (1982). Constructive logic versus algebraization [I]. In A. S. Troelstra & D. van Dalen (Eds.), *The L.E.J. Brouwer centenary symposium* (pp. 217–260). North-Holland.
Kreisel, G., & Newman, M. (1969). Luitzen Egbertus Jan Brouwer. 1881–1966. *Biographical Memoirs of Fellows of the Royal Society*, *15*, 39–68.
Kreisel, G., & Takeuti, G. (1974). Formally self-referential propositions for cut free classical analysis and related systems. *Polska Akademia Nauk. Instytut Matematyczny. Dissertationes Mathematicae. Rozprawy Matematyczne*, *118*, 55.
Kreisel Papers. SC0136, Department of Special Collections, Stanford University Archives. Cited by box/folder.
Kuroda, S. (1956). Review of *Intuitionism. An Introduction*, by Arend Heyting. *Journal of Symbolic Logic*, *21*(4), 367–371.
Lakatos, I. (1963–1964). Proofs and refutations. *British Journal for the Philosophy of Science*, *14*(53), 1–25, 120–139, 221–245, 296–342.
Lawvere, F. W. (1969). Diagonal arguments and cartesian closed categories. In *Category theory, homology theory and their applications, II: Battelle Institute Conference, Seattle, Wash., 1968. Vol. 2* (pp. 134–145). Springer, Berlin.
Leitgeb, H. (2009). On formal and informal provability. In O. Bueno & Ø. Linnebo (Eds.), *New waves in philosophy of mathematics* (pp. 263–299). Palgrave Macmillan.
Li, X. (2021). Answer to 'Y combinator as an application of Lawvere's fixed point theorem'. https://math.stackexchange.com/questions/2823016/y-combinator-as-an-application-of-lawveres-fixed-point-theorem
Löb, M. (1955). Solution of a problem of Leon Henkin. *Journal of Symbolic Logic*, *20*, 115–118.
Lombard, M., & Vesley, R. (1998). A common axiom set for classical and intuitionistic plane geometry. *Annals of Pure and Applied Logic*, *95*(1-3), 229–255.
Longa, G. (2020). 'The hidden analysis': On the spread of the analytic method in Greek geometry [Slides for a presentation at 'Research in Progress', British Society for the History of Mathematics, Oxford, February 29, 2020.].
Mäenpää, P. (1993). *The art of analysis* (Doctoral dissertation). University of Helsinki. Helsinki.
Martin-Löf, P. (2006). 100 years of Zermelo's axiom of choice: what was the problem with it? *The Computer Journal*, *49*(3), 345–350.
Martin-Löf, P. (2008). The Hilbert-Brouwer controversy resolved? In M. van Atten, P. Boldini, M. Bourdeau, & G. Heinzmann (Eds.), *One hundred years of Intuitionism (1907–2007): The Cerisy conference* (pp. 243–256). Birkhäuser.
Mauro, L., & Venturi, G. (2015). Naturalness in mathematics. In G. Lolli, M. Panza, & G. Venturi (Eds.), *From logic to practice: Italian studies in the philosophy of mathematics* (pp. 277–313). Springer.
Menn, S. (2002). Plato and the method of analysis. *Phronesis*, *47*(3), 192–223.
Morris, R. (2020). Motivated proofs: What they are, why they matter and how to write them. *Review of Symbolic Logic*, *13*(1), 23–46.
Myhill, J. (1960). Some remarks on the notion of proof. *Journal of Philosophy*, *57*, 461–471.

Netz, R. (2000). Why did Greek mathematicians publish their analyses? In *Ancient & medieval traditions in the exact sciences (Stanford, CA, 1998)* (pp. 139–157). CSLI Publ., Stanford, CA.
Odifreddi, P. (2019a). In *About logic and logicians: A palimpsest of essays by Georg Kreisel. Selected and arranged by Piergiorgio Odifreddi. Vol. 1. Philosophy*. Lógica no Avião.
Odifreddi, P. (2019b). In *About logic and logicians: A palimpsest of essays by Georg Kreisel. Selected and arranged by Piergiorgio Odifreddi. Vol. 2. Mathematics*. Lógica no Avião.
Pappus of Alexandria. (1986). *Book 7 of the Collection* (A. Jones, Ed. & Trans.). Springer.
Parsons, C. (2020). Kreisel and Gödel. In P. Weingartner & H.-P. Leeb (Eds.), *Kreisel's interests* (pp. 63–85). College Publications.
Pascal, B. (1936). De l'esprit géométrique et de l'art de persuader. In J. Chevalier (Ed.), *L'Œuvre de Pascal* (pp. 358–386). Gallimard. (Original work published 1658)
Petrakis, I. (n.d.). Review of Weaver (2012). *Zentralblatt für Mathematik und ihre Grenzgebiete*, *Zbl 1264.03031*.
Placek, T. (1999). *Mathematical intuitionism and intersubjectivity: A critical exposition of arguments for intuitionism*. Kluwer.
Poincaré, H. (1906). Les mathématiques et la logique. *Revue de métaphysique et de morale*, (14), 294–317.
Pólya, G. (2004). *How to solve it: A new aspect of mathematical method* [First edition 1945]. Princeton University Press.
Prawitz, D. (1970). Some results for intuitionistic logic with second order quantification rules. In A. Kino, J. Myhill, & R. Vesley (Eds.), *Intuitionism and proof theory: Proceedings of the summer conference at Buffalo N. Y., 1968* (pp. 259–269). North-Holland.
Prawitz, D. (1971). Ideas and results in proof theory. In J. Fenstad (Ed.), *Proceedings of the second Scandinavian logic symposium* (pp. 235–307). North-Holland.
Presburger, M. (1930). Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt. In *Sprawozdanie z i kongresu matematyków krajów slowiańskich, warszawa 1929* (pp. 92–101, 395).
Ramsey, F. (1931). *The foundations of mathematics and other logical essays*. Routledge & Kegan Paul.
Rasiowa, H., & Sikorski, R. (1963). *The mathematics of metamathematics*. Państwowe Wydawnictwo Naukowe, Warsaw.
Roberts, D. M. (2021). Substructural fixed-point theorems and the diagonal argument: Theme and variations. https://arxiv.org/abs/2110.00239
Rogerson, S. (2007). Natural deduction and Curry's paradox. *Journal of Philosophical Logic*, *36*, 155–179.
Rood, P. T. M. (2005). *Proofs, intuitions and diagrams: Kant and the mathematical method of proof* (Doctoral dissertation). Vrije Universiteit Amsterdam. Amsterdam.
Ruitenburg, W. (1991). The unintended interpretations of intuitionistic logic. In T. Drucker (Ed.), *Perspectives on the history of mathematical logic* (pp. 134–160). Birkhäuser Boston, Boston, MA.
Runciman, W. G. (1962). *Plato's later epistemology*. Cambridge University Press.
Russell, B. (1903). *The principles of mathematics*. Allen & Unwin.
Russell, B. (1906). On some difficulties in the theory of transfinite numbers and order types. *Proc. London Mathematical Society*, *4*, 29–53.
Santos, P. (2020). *Diagonalization in formal mathematics*. Springer.
Schilpp, P. (Ed.). (1944). *The Philosophy of Bertrand Russell*. Nortwestern University Press.
Scott, D. (1970). Constructive validity. In M. Laudet, D. Lacombe, L. Nolin, & M. Schützenberger (Eds.), *Symposium on automatic demonstration, Versailles, December 1968* (237–275). Springer.
Smoryński, C. (1991). The development of self-reference: Löb's theorem. In T. Drucker (Ed.), *Perspectives on the history of mathematical logic* (pp. 110–133). Birkhäuser Boston, Boston, MA.
Smullyan, R. (1965). Analytic natural deduction. *Journal of Symbolic Logic*, *30*, 123–139.

Smullyan, R. (1968). Analytic cut. *Journal of Symbolic Logic*, *33*, 560–564.

Smullyan, R. (1978). *What is the name of this book? The riddle of dracula and other logical puzzles*. Prentice-Hall.

Spector, C. (1962). Provably recursive functionals of analysis: A consistency proof of analysis by an extension of principles formulated in current intuitionistic mathematics. In *Recursive function theory* (1–27).

Straßburger, L. (2019). The problem of proof identity, and why computer scientists should care about Hilbert's 24th problem. *Philosophical Transactions of the Royal Society of London. A. Mathematical, Physical and Engineering Sciences*, *377*(2140).

Sundholm, G. (1983). Constructions, proofs and the meaning of logical constants. *Journal of Philosophical Logic*, *12*, 151–172.

Sundholm, G. (2004). The proof-explanation of logical constants is logically neutral. *Revue internationale de philosophie*, (4), 401–410.

Sundholm, G., & van Atten, M. (2008). The proper interpretation of intuitionistic logic: On Brouwer's demonstration of the Bar Theorem. In M. van Atten, P. Boldini, M. Bourdeau, & G. Heinzmann (Eds.), *One hundred years of Intuitionism (1907–2007): The Cerisy conference* (pp. 60–77). Birkhäuser.

Tanswell, F. (2015). A problem with the dependence of informal proofs on formal proofs. *Philosophia Mathematica*, *23*(3), 295–310.

Tarski, A. (1948). *A decision method for elementary algebra and geometry* [Prepared for publication by J. C. C. McKinsey]. The Rand Corporation.

Troelstra, A. S. (1969). *Principles of intuitionism*. Springer.

Troelstra, A. S. (Ed.). (1973a). *Metamathematical investigation of intuitionistic arithmetic and analysis*. Springer.

Troelstra, A. S. (1973b). Notes on intuitionistic second order arithmetic. In H. Rogers & A. Mathias (Eds.), *Cambridge summer school in mathematical logic 1971* (pp. 171–205). Springer.

Troelstra, A. S. (1982). Intuitionistic extensions of the reals [II]. In *Logic Colloquium '80 (Prague, 1980)* (pp. 279–310). North-Holland.

Troelstra, A. S. (2018). Corrections to some publications. https://eprints.illc.uva.nl/id/eprint/1650/1/CombiCorr101218.pdf

Troelstra, A. S., & van Dalen, D. (1988). *Constructivism in mathematics* [2 vols.]. North-Holland.

van Atten, M. (2004). *On Brouwer*. Wadsworth.

van Atten, M. (2009). The hypothetical judgement in the history of intuitionistic logic: Proceedings of the 13th international congress, Bejing 2007. In C. Glymour, W. Wang, & D. Westerstahl (Eds.), *Logic, methodology, and philosophy of science* (122–136). College Publications.

van Atten, M. (2015). *Essays on Gödel's reception of Leibniz, Husserl, and Brouwer*. Springer.

van Atten, M. (2017a). The development of intuitionistic logic. In E. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2017). Metaphysics Research Lab, Stanford University. http://plato.stanford.edu/archives/win2017/entries/intuitionistic-logic-development/

van Atten, M. (2017b). Predicativity and parametric polymorphism of Brouwerian implication. https://hal.archives-ouvertes.fr/halshs-01619110

van Atten, M. (2017c). Troelstra's Paradox and Markov's Principle. https://hal.archives-ouvertes.fr/hal-01415603

van Atten, M. (2018). The Brouwer-Kripke Schema, the Creating Subject, and infinite proofs. *Indagationes Mathematicae*, *29*, 1565–1636.

van Atten, M., & Sundholm, G. (2017). L.E.J. Brouwer's 'Unreliability of the logical principles': A new translation, with an introduction. *History and Philosophy of Logic*, *38*(1), 24–47.

van Benthem, J. (1978). Four paradoxes. *Journal of Philosophical Logic*, *7*, 49–72.

van Dalen, D. (2004). *Logic and Structure* (4th ed.). Springer.

van Heijenoort, J. (Ed.). (1967). *A sourcebook in mathematical logic, 1879–1931*. Harvard University Press.

von Plato, J. (2017). *The great formal machinery works: Theories of deduction and computation at the origins of the digital age*. Princeton University Press.

Wadler, P. (2014). Propositions as types. https://homepages.inf.ed.ac.uk/wadler/papers/propositions-as-types/propositions-as-types.pdf

Wang, H. (1981). Some facts about Kurt Gödel. *Journal of Symbolic Logic*, *46*(3), 653–659.

Wang, H. (1987). *Reflections on Kurt Gödel*. MIT Press.

Wang, H. (1996). *A logical journey: From Gödel to philosophy*. MIT Press.

Warren, L., Diener, H., & McKubre-Jordens, M. (2018). The Drinker Paradox and its dual. https://arxiv.org/abs/1805.06216

Weaver, N. (2012). Intuitionism and the liar paradox. *Annals of Pure and Applied Logic*, *163*(10), 1437–1445.

Weinstein, S. (1983). The intended interpretation of intuitionistic logic. *Journal of Philosophical Logic*, *12*, 261–270.

Weyl, H. (1921). Über die neue Grundlagenkrise der Mathematik. *Mathematische Zeitschrift*, *10*, 39–79.

Weyl, H. (1928). Diskussionsbemerkungen zu dem zweiten Hilbertschen Vortrag über die Grundlagen der Mathematik. *Abhandlungen des mathematischen Seminars zu Hamburg*, *6*, 86–88.

Weyl, H. (1946). Review of Schilpp (1944). *The American Mathematical Monthly*, *53*(4), 208–214.

Yanofsky, N. (2003). A universal approach to self-referential paradoxes, incompleteness and fixed points. *Bulletin of Symbolic Logic*, *9*(3), 362–386.

Index of references to equations and quotations