1–36
10.1093/philmat/nkaa010
Philosophia Mathematica

Does Choice Really Imply Excluded Middle? Part II: Historical, Philosophical, and Foundational Reflections on the Goodman–Myhill Result†

Neil Tennant*

Department of Philosophy, The Ohio State University, Columbus, Ohio 43210, U.S.A.

ABSTRACT

Our regimentation of Goodman and Myhill's proof of Excluded Middle revealed among its premises a form of Choice and an instance of Separation.

Here we revisit Zermelo's requirement that the separating property be *definite*. The instance that Goodman and Myhill used is not constructively warranted. It is *that* principle, and not Choice alone, that precipitates Excluded Middle.

Separation in various axiomatizations of constructive set theory is examined. We conclude that insufficient critical attention has been paid to how those forms of Separation fail, in light of the Goodman–Myhill result, to capture a genuinely constructive notion of set.

1. A CRITIQUE OF GOODMAN AND MYHILL'S ABSTRACTION PRINCIPLE

Let us think classically and rather naïvely about 'the set' whose existence is asserted by the principle

$$\exists y \forall x (x \in y \leftrightarrow (x = 0 \wedge \psi)).$$

†Thanks for helpful comments on earlier drafts are owed to Mark van Atten, John Bell, Ethan Brauer, Laura Crosilla, David Charles McCarty, Michael Potter, Michael Rathjen, Giorgio Sbardolini, Stewart Shapiro, and William Tait. The author is especially grateful to anonymous referees for *Philosophia Mathematica* for insightful criticisms.

*Orcid.org/0000-0002-3523-2296. E-mail: tennant.9@osu.edu

The constructivist should use scare quotes even with certain 'small-looking' sets, for reasons that will emerge below.

Let t be such that

$$\forall x(x \in t \leftrightarrow (x{=}0 \wedge \psi)).$$

This biconditional can be unpacked as assertable on the basis of two proofs of the following respective forms (each of them parametric in a):

$$\begin{array}{ccc} a{=}0 \wedge \psi & \qquad & a \in t \\ \vdots & & \vdots \\ a \in t & & a{=}0 \wedge \psi \end{array}$$

Note that the assumption of each proof implies $\exists! a$; so one need not be detained by any free-logic worries on that score.

One might be (classically and naïvely) tempted to think that t, under these conditions, is 'really' the set of all x such that $x{=}0 \wedge \psi$. So, this line of thought would go, one ought to be able to establish the formal conclusion

$$t = \{x|x{=}0 \wedge \psi\}.$$

It would appear, moreover, that t can take only one of two forms — that is, its denotation can *be* only one of two possible sets. Suppose on one hand that ψ is true. Then (surely)

$$t = \{0\}.$$

Suppose on the other hand that ψ is false. Then (surely) t is empty:

$$t = \emptyset,$$

since the defining predicate $x = 0 \wedge \psi$ will not be satisfied by *anything* — not even by 0. And that is pretty innocuous, this classical thinker will think. After all, even the constructivist has the axiom stating that the empty set exists; and has also the pair-set axiom, which gives us singletons of any existents for free. The set t is not problematically infinite for the constructivist like some very large cardinal, or even, 'lower down', the countably infinite set ω of all finite ordinals. So — surely? — the set t (*i.e.*, $\{x|x{=}0 \wedge \psi\}$) is either the empty set or singleton 0.

This line of thinking is erroneous. The thinker here is exploiting the strictly classical form of reasoning known as Dilemma. This involves supposing on the one hand that ψ is true, and deriving the sought conclusion (call it θ); then supposing on the other hand that ψ is false, and deriving the same conclusion; and, finally, discharging the horn-assumptions ψ and $\neg\psi$ in the respective subordinate proofs:

$$\dfrac{\begin{array}{cc} {}^{(i)}\overline{\psi} & \overline{\neg\psi}^{(i)} \\ \vdots & \vdots \\ \theta & \theta \end{array}}{\theta}{\scriptstyle (i)}$$

This way of reasoning to the existence of $\{x|x = 0 \wedge \psi\}$ (as the sought conclusion θ) is strictly classical. (It can be regarded as a proof-by-cases whose major premise is $\psi \vee \neg\psi$). It presupposes that the world is determinate in the ψ-regard. For undecidable, or even undecided, sentences ψ it affords the thinker no grasp at all of what, exactly, 'the set' in question *is*. What exactly are its members? The constructivist does not wish to be told 'Oh, only 0, provided ψ is true; otherwise, it has no members at all.'

Let us revisit the rule $\{\ \}$I in order to see how it lends support to the constructivist's case here. Let the formula $x = 0 \wedge \psi$ play the role of Φ in the statement of that rule. One would need to supply the three subordinate proofs indicated:

$$\{\ \}\mathrm{I}: \qquad \dfrac{\begin{array}{ccc} \underbrace{{}^{(i)}\overline{\exists! a}\ ,\ \overline{\Phi^x_a}^{(i)}} & & \overline{a \in t}^{(i)} \\ \vdots & \vdots & \vdots \\ a \in t & \exists! t & \Phi^x_a \end{array}}{t = \{x|\Phi\}}{\scriptstyle (i)} \quad , \quad \text{where } a \text{ is parametric}$$

Fully spelled out with $x = 0 \wedge \psi$ playing the role of Φ, this would take the form

$$\dfrac{\begin{array}{ccc} \underbrace{{}^{(i)}\overline{\exists! a}\ ,\ \overline{a = 0 \wedge \psi}^{(i)}} & & \overline{a \in t}^{(i)} \\ \vdots & \vdots & \vdots \\ a \in t & \exists! t & a = 0 \wedge \psi \end{array}}{t = \{x|x = 0 \wedge \psi\}}{\scriptstyle (i)}$$

Note that the mention (and discharge) of the assumption $\exists! a$ at top left is unnecessary, since the other assumption in that subordinate proof implies it:

$$\dfrac{\dfrac{a = 0 \wedge \psi}{a = 0}}{\exists! a}$$

So the required form of { }-Introduction for the sought conclusion simplifies to

$$\dfrac{\begin{array}{c}\overline{a=0\wedge\psi}^{(i)}\\ \vdots\\ a\in t\end{array}\qquad\begin{array}{c}\\ \vdots\\ \exists!t\end{array}\qquad\begin{array}{c}\overline{a\in t}^{(i)}\\ \vdots\\ a=0\wedge\psi\end{array}}{t=\{x|x=0\wedge\psi\}}{\scriptstyle(i)}$$

And *this* looks familiar. It is the inferential unpacking of the biconditional that was remarked on earlier, but with some added extra. That added extra is an important fly in the ointment: the middle premise $\exists!t$. This requires the thinker *already to have established the existence of* t, or to be simply assuming it. This is the intellectual discipline that free logic imposes on one's reasoning about sets.

What Goodman and Myhill really needed (and simply helped themselves to) is the following form of the rule, stripped of this irksome existential requirement on t:

$$\dfrac{\begin{array}{c}\overline{a=0\wedge\psi}^{(i)}\\ \vdots\\ a\in t\end{array}\qquad\begin{array}{c}\overline{a\in t}^{(i)}\\ \vdots\\ a=0\wedge\psi\end{array}}{t=\{x|x=0\wedge\psi\}}{\scriptstyle(i)}\quad .$$

The constructivist cannot, however, allow them to engage in such premise suppression.

Recall that the rules of introduction and elimination supply only conceptual constraints governing the interplay of set abstraction, predication and membership. They do *not* involve, incur, or presuppose the existence of any set at all. So the constructive logicist who is considering the abstraction principle

$$\exists y\forall x(x\in y\leftrightarrow(x=0\wedge\psi))$$

is perfectly entitled to object that it carries unwarrantable ontological (or metaphysical) commitment. It matters not that the indeterminacy of the set t that the principle is supposed to afford arises from 'merely vacillating' between a particular singleton and the empty set, depending on how the world is in the ψ-regard. The belief that the foregoing abstraction principle is true is based on strictly classical reasoning which, from the constructivist's point of view, furnishes no warrant at all for the existence of 'the set' in question. The worry transfers, as was foreshadowed in the Introduction to Part I, to the closely related (and equivalent)[1] abstraction principle

$$\exists y\forall x(x\in y\leftrightarrow x\in\{0\}\wedge\psi),$$

[1] The equivalence of the two abstraction principles follows from the theorem that

$$\forall x\forall y(x\in\{y\}\leftrightarrow x=y).$$

which is but a special case of Bell's axiom of **WST** that he called **Restricted Subsets**:[2]

$$\exists y \forall x(x \in y \leftrightarrow x \in a \wedge \psi).$$

This is in agreement with Tait, who, though he was not directly considering the Goodman–Myhill result or the structure of the reasoning therein, expressed the following insight:

> It is not applications of AC that lead to existence proofs which give no means of identifying an object of the kind proved to exist. Rather, it is applications of LEM. [1994, p. 65, n. 8]

A set-abstraction term's failure to denote can result from various sources. Perhaps the predicate involved has too large an extension, such as $x = x$. Perhaps it is internally contradictory, given other basic principles of set existence, as is the case with the predicate $\neg\, x \in x$. Perhaps its extension is indefinitely extensible, as with 'x is an ordinal'. And finally: *perhaps the defining predicate's holding, or not holding, of any given object depends on the world being determinate in certain regards, which one is not justified in assuming.*

The basic lesson to draw from the ingenious piece of reasoning furnished by Goodman and Myhill is not that (full) Choice implies Excluded Middle. It is rather that an innocuous-*looking* but surreptitiously and objectionably *realist* set-existence principle turns out to have been adopted.

The introduction and elimination rules that were given above for set formation have been framed within a free logic so as to be ontologically noncommittal. This has the virtue that, even though the resulting notion of set is fully *extensional*, one can nevertheless find fault, in a principled way, with the Goodman–Myhill attempt to have Choice visit Excluded Middle upon the previously unsuspecting constructive set theorist. It is not Choice that is doing the unwelcome work; it is too lax an admission of certain sets as existing.

If one decides, consequently, to place the blame on the particular kind of *set-existence* principles involved,[3] then one might have to reconsider the (now standard) inclusion, in intuitionistic or constructive set theory, of the Axiom Scheme of Separation, which does not (even) require that ψ be a restricted formula:

$$\exists y \forall x(x \in y \leftrightarrow x \in a \wedge \psi).$$

For, if one blames this principle in situations where ψ *is* a restricted formula, then one can hardly espouse Separation in the more general form.

---

[2] See § 4.2.1 of Part I for discussion of Bell's axioms.

[3] Note that one can so re-apportion the blame even while retaining the ontologically non-committal introduction and elimination rules for set abstraction.

1.1. Zermelo 1908

Let us now revisit the seminal article [Zermelo, 1908] in order to examine his founding discussion of Separation. Reason will be uncovered to think that constructivists and intuitionists could — in response to the Goodman–Myhill result — take one particular lead that Zermelo happened to offer there. It would of course be highly anachronistic to think that Zermelo himself might have had anything like that looming problem in mind, on behalf of the intuitionist, or would have been at all interested in providing some sort of argument on behalf of intuitionists by way of a helpful suggestion as to how they might frame their response. But the tradition seems largely to have ignored Zermelo's own careful remarks about the applicability of Separation. Set theorists have simply taken the formal statement of Separation as an axiom scheme of universal applicability within their first-order language. And they have done so within the framework of classical logic, premised by default on the belief that every declarative sentence has a determinate truth value, and that every monadic formula determinately holds, or fails to hold, of any given object in the domain. It is precisely these default assumptions that Zermelo himself can be read as questioning (or at least as preparing the ground for a particular line of critical questioning).

Zermelo deployed the important metaconcept of *definiteness* (on the part of a proposition or concept expressed in the formal language of set theory). Here is how he introduced the notion:[4]

> 4. A question or assertion $\mathfrak{E}$ is is said to be *definite* if the fundamental relations of the domain, by means of the axioms and the universally valid laws of logic, determine without arbitrariness whether it holds or not. Likewise a "propositional function" ⟦"Klassenaussage"⟧ $\mathfrak{E}(x)$, in which the variable x ranges over all individuals of a class $\mathfrak{K}$, is said to be definite if it is definite for *each single* individual x of the class $\mathfrak{K}$.

Zermelo then proceeded to state his axiom scheme of Separation as follows:[5]

---

[4] [van Heijenoort, 1967, p. 201]. The original German [Zermelo, 1908, p. 263], reads as follows.

> Eine Frage oder Aussage $\mathfrak{E}$, über deren Gültigkeit oder Ungültigkeit die Grundbeziehungen des Bereiches vermöge der Axiome und der allgemeingültigen logischen Gesetze ohne Willkür entscheiden, heißt „*definit*". Ebenso wird auch eine ,,Klassenaussage" $\mathfrak{E}(x)$, in welcher der variable Term x alle Individuen einer Klasse $\mathfrak{K}$ durchlaufen kann, als „definit" bezeichnet, wenn sie für *jedes einzelne* Individuum x der Klasse $\mathfrak{K}$ definit ist.

[5] *Ibid.*, p. 202. The original German [Zermelo, 1908, p. 263] reads as follows:

> Axiom III. Ist die Klassenaussage $\mathfrak{E}(x)$ definit für alle Element einer Menge M, so besitzt M immer eine Untermenge $M_{\mathfrak{E}}$, welche alle diejenigen Elemente x von M, für welche $\mathfrak{E}(x)$ wahr ist, und nur solche Element enthält.

> Axiom III (Axiom of separation ⟦Axiom der Aussonderung⟧.) Whenever the propositional function $\mathfrak{E}(x)$ is definite for all elements of a set M, M possesses a subset $M_{\mathfrak{E}}$ containing as elements precisely those elements x of M for which $\mathfrak{E}(x)$ is true.

In formalizing this informally rigorous formulation of Zermelo, one has to ask oneself whether it is faithful to his original intent simply to say (as set theorists do nowadays) that the theory has as axioms all instances of the following scheme, where $\varphi(x)$ is a formula in the language of set theory (whose primitive predicates are $\in$ and $=$) that has x as its only free variable:

$$\forall z \exists y \forall x (x \in y \leftrightarrow (x \in z \wedge \varphi(x)))\,;$$

— or, if the variable-binding set-abstraction operator is available as a primitive:

$$\forall z \exists ! \{x | x \in z \wedge \varphi(x)\}\,;$$

— or, to express the same principle in rule form:[6]

$$\frac{\exists ! t}{\exists ! \{x \in t | \varphi\}}\ .$$

For this is to suppress altogether Zermelo's apparent concern about the *Definitheit* (for all elements of the given set x) of the 'separating formula' φ. Only a classical logician could think that φ (provided it is in the language of set theory) will be definite *by default.* But in that case one would have to interpret as completely otiose, or beside the point, Zermelo's very express concern about *Definitheit.*

1.2. First Historical Interlude: Dedekind

Zermelo would certainly have been familiar with the following passage from [Dedekind, 1888, §1.2],[7]

---

[6] Here t is a closed term. The familiar abbreviation $\{x \in t | \varphi\}$ is also used, for the set term $\{x | x \in t \wedge \varphi\}$.

[7] Thanks are owed to an anonymous referee for bringing this passage from Dedekind to the author's attention. The German reads as follows:

> 2. Es kommt sehr häufig vor, daß verschiedene Dinge a, b, c ... aus irgendeiner Veranlassung unter einem gemeinsamen Gesichtspunkte aufgefasst, im Geiste zusammengestellt werden, und man sagt dann, daß sie eine **Menge** S bilden; man nennt die Dinge a, b, c ... die **Elemente** der Menge S, sie sind **enthalten** in S; umgekehrt **besteht** S aus diesen Elementen. Eine solche Menge S (oder ein Inbegriff, eine Mannigfaltigkeit, eine Gesamtheit) ist als Gegenstand unseres Denkens ebenfalls ein Ding ...; es ist vollständig bestimmt, wenn von jedem Ding bestimmt ist, ob es ein Element von S ist oder nicht [5]).

> It often happens that different things a, b, c ... grasped for some reason from a common standpoint, are brought together in the mind, and one then says that they form a **set** S; one calls the things a, b, c ... the **elements** of the set S, that they are **contained** in S; conversely, S **consists** of these elements. Such a set S (or a concept, a manifold, a totality) is as an object of our thought likewise a thing ...; it is completely determined, if of each thing it is determined whether it is an element of S or not [5]).

Footnote 5 reads:

> In what way this determination comes about, and whether we know a way to decide such matters, does not matter for anything that follows; the general laws to be developed do not depend on this, they hold under all circumstances. I mention this explicitly because recently Mr. **Kronecker** (*Journal für Mathematik* **99**, 334–336) has wished to impose certain restrictions on free concept-formation in mathematics, which I do not acknowledge as justified; but to go into this more closely appears to be warranted only when this excellent mathematician will have published his grounds for the necessity or even just the advisability of these restrictions.

Kronecker's misgivings, in the work referenced by Dedekind [Kronecker, 1886], concerned the passage from certain finite systems to infinite ones of a supposedly similar kind; the exact details of his proposed restrictions need not concern us here.

It is the usual practice nowadays to have Separation (and other axiom schemes) expressed in terms of formulae in the language of set theory based on $\in$ and $=$ as the sole primitive predicates. It is worth pointing out, however, that when Zermelo wrote of 'the fundamental relations of the domain' it was a clarifying advance so to limit the stock of primitive predicates. For otherwise it would have been possible for his reader to consider including such predicates as 'Dedekind has thought of x'.[8]

Footnote 5 reads:

> Auf welche Weise diese Bestimmtheit zustande kommt, und ob wir einen Weg kennen, um hierüber zu entscheiden, ist für alles Folgende gleichgültig; die zu entwickelnden allgemeinen Gesetze hängen davon gar night ab, sie gelten unter allen Umständen. Ich erwähne dies ausdrücklich, weil Herr **Kronecker** vor kurzem (in Band 99 des Journals für Mathematik, S. 334–336) der freien Begriffsbildung in der Mathematik gewisse Beschränkungen hat auferlegen wollen, die ich nicht als berechtigt anerkenne; näher hierauf einzugehen erscheint aber erst dann geboten, wenn der ausgezeichnete Mathematiker seine Gründe für die Notwendigkeit oder auch nur die Zweckmässigkeit dieser Beschränkungen veröffentlicht haben wird.

[8] See [Dedekind, 1888, Theorem 66]. From the rather unhappy English translation [Dedekind, 1901, p. 64]:

It should also be borne in mind that Zermelo was writing at a time when there was not yet any clear distinction in logicians' minds between logical consequence and deducibility, nor any appreciation of the fact that axiomatic theories can have models in which certain truths lie beyond the reach of proof *and* beyond the reach of consequence (from the axioms).

In this connection it is interesting that Zermelo added quite a lengthy comment on his Axiom III, which in the German original (but not in the van Heijenoort source book of English translations) appears in a smaller font, suggesting that it might well have been added as an important afterthought closer to the time of publication than the original submission. In this extended comment Zermelo saw fit to re-emphasize the following:[9]

> ...the defining criterion [the separating formula — Author] must always be definite in the sense of our definition in No. 4 (that is, for each single element x of M the fundamental relations of the domain must determine whether it holds or not). ... *we must, prior to each application of our* Axiom III*, **prove** the criterion $\mathfrak{E}(x)$ in question to be definite, if we wish to be rigorous* ... [Emphasis and double emphasis added.]

In light of the importance Zermelo accords to his requirement of *Definitheit* of the separating property, it is exegetically reasonable to suggest that the rule form given above for Separation:

$$\frac{\exists! t}{\exists! \{x \in t | \varphi\}}$$

66. Theorem. There exist infinite systems.

Proof.[fn] My own realm of thoughts, *i.e.*, the totality S of all things, which can be objects of my thought, is infinite. For if s signifies an element of S, then is [*sic*] the thought s', that s can be object [*sic*] of my thought, itself an element of S. If we regard this as transform [*sic*] $\phi(s)$ of the element s then has [*sic*] the transformation ϕ of S, thus determined, the property that the transform S' is part of S; and S' is certainly proper [*sic*] part of S, because there are elements in S (*e.g.*, my own ego) which are different from such [*sic*] thought s' and therefore are not contained in S'. Finally it is clear that if a, b are different elements of S, their transforms a', b' are also different, that therefore the transformation ϕ is a distinct (similar) transformation ... . Hence S is infinite, which was to be proved.

[9] *Ibid.*, p. 202. The original German, [Zermelo, 1908, p. 264], reads as follows.

Zugleich muss ... das bestimmte Kriterium $\mathfrak{E}(x)$ im Sinne unserer Erklärung Nr. 4 immer „*definit*" d. h. für jedes einzelne Element x von M durch die „Grundbeziehungen des Bereiches" entschieden sein ... Hieraus folgt aber auch daß, streng genommen, vor jeder Anwendung unseres Axiom III immer erst das betreffende Kriterium $\mathfrak{E}(x)$ als „*definit*" nachgewiesen werden muß ...

is actually sorely lacking. For it does not incorporate the requirement that the separating formula φ be definite. The question is: how best to incorporate it? A certain degree of *explication* will be called for.

One would not be alone, or original, in wondering how. In [Fraenkel, 1922, pp. 231–232] there is the following criticism of Zermelo's definition of *definit*:[10]

> The one weak point in Zermelo's axiomatization is the 'Definition' of the concept of a 'definite' question or proposition about classes [[Zermelo, 1908, No. 4 and Axiom III, p. 263]]. The Zermeloan standpoint (following that of Cantor) which here may well stand in the sharpest contrast with the Brouwerian conception, requires within the scope of his theory a sharp precisification, which proceeds appropriately from the assumption that the relations $m\varepsilon M$ ("m is an element of M") and $M = N$ are declared to be definite.

Fraenkel expresses unease once again over the unclarity of Zermelo's crucial notion of *Definitheit*. He writes (at p. 234):[11]

> ... the insufficient determinacy of the concept *definite* [is] very disturbing ...

and later (on p. 236) Fraenkel proposes the following *determinate* sense ('*bestimmter* Sinn') for the concept ,,Definit":[12]

> ... for example one may speak of a definite proposition $\mathfrak{E}$ if it can be formulated by means of the basic relations $m\varepsilon M$ and $M = N$ "in the sense

[10] English translation by the author. The original German reads

> ... der einzige schwache Punkt in Zermelos Axiomatik ist die ,,Definition" des Begriffs einer ,,*definiten*" Frage oder Klassenaussage ([Zermelo, 1908], Nr. 4 und Axiom III, S. 263). Der Zermelosche (an Cantor sich anschließende) Standpunkt, der hier wohl in schärfstem Gegensatz zur Brouwerschen Auffassung steht, bedarf im Rahmen seiner Theorie einer scharfen Präzisierung, die zweckmäßigerweise davon ausgeht, daß von vornherein die Beziehungen $m\varepsilon M$ (,,m ist Element von M") und $M = N$ als definit erklärt werden.

[11] English translation by the author. The original German reads

> ... die ungenügende Bestimmtheit des Begriffs ,,Definit" [ist] sehr störend ... .

[12] English translation by the author. The original German reads

> ... z.B. mag von einer definiten Aussage $\mathfrak{E}$ gesprochen werden, falls sie mittels der Grundbeziehungen $m\varepsilon M$ und $M = N$,,im Sinn des heutigen Standes der Wissenschaft" formuliert *und ,,intern entscheidbar"* gedacht werden kann. [Emphasis added.]

of the current condition of science", *and can be thought of as "internally decidable"*. [Emphasis added.]

In light of this one could contend that the rule form for Separation, *for the constructivist*, could and should be expressed as follows:

$$\dfrac{\exists! t \qquad \begin{array}{c}\overline{a \in t}^{(i)} \\ \vdots \\ \varphi(a) \vee \neg\varphi(a)\end{array}}{\exists!\{x \in t | \varphi(x)\}}{\scriptstyle (i)} \qquad a \text{ parametric.}$$

The right-hand subproof codifies, inferentially, the requirement that φ be (proved to be) definite on every individual in t.[13] If one were instead to prefer a sentential, axiomatic form for Separation, the appropriately modified scheme would be

$$\forall z(\forall x(x \in z \rightarrow (\varphi(x) \vee \neg\varphi(x))) \rightarrow \exists!\{x | x \in z \wedge \varphi(x)\}).$$

One stresses '*for the constructivist*' because, in the form of the rule currently suggested, the subordinate conclusion $\varphi(a) \vee \neg\varphi(a)$ would *not*, for the constructivist, be trivially provable, as it would be for the classicist.

It is this trivial provability that explains why, within the classical tradition, Zermelo's requirement of *Definitheit* has simply withered on the vine. This has perhaps resulted in those constructivists visiting set-theoretic foundations much later in the 1970s, omitting to consider Zermelo's original requirement and the potential wisdom of re-imposing it.

In its current form the rule (or axiom) being suggested here has heft only for the constructivist. For the classicist, by contrast, the 'definiteness' condition on $\varphi(x)$ can be met by default (the advantage of classical theft over honest constructivist toil). If Zermelo himself had (as is likely to have been the case) no scruples about using classical inferences, he would have been looking for a *different* way to capture the definiteness condition on $\varphi(x)$ (as a premise in the rule form, or as an extra conjunct in the antecedent of the corresponding conditional form). To do so, however, he would in all likelihood have had to ascend to the metalevel, with a rule of the following sort of form (where, again, a is parametric):

$$\dfrac{\exists! t \quad \begin{array}{c}\overline{a \in t}^{(i)} \\ \vdots \\ \text{for some term } u \text{ denoting } a,\ \vdash \varphi(u)\ \mathit{OR} \text{ for some term } u \text{ denoting } a,\ \varphi(u) \vdash \bot\end{array}}{\exists!\{x \in t | \varphi\}}{\scriptstyle (i)}$$

[13] There will be a fuller discussion of the adequacy of the rule form of Separation being proposed here, when it is revisited in §3.

With the requirement of definiteness formulated in this way (which, admittedly, is rather clumsy) one can at least see that deploying classical logic will not lead to its automatic fulfillment, even if the embedded single turnstiles were to be read as deducibility in *classical* logic.

This attempt to capture what it would be to make the requirement of definiteness non-trivial for the classicist suffers, though, from the drawback that the axiom (or rule) scheme of Separation would no longer be one whose instances are all *in the object language* of set theory. This is not to say, however, that something like this rule form requirement would not have crossed Zermelo's mind. He was, after all, writing in the Göttingen of 1907–1908, where Hilbert had recently axiomatized (first-order) Euclidean geometry by resort to what was, in effect, a similarly *meta* linguistic Axiom of Completeness. The distinction between object language and metalanguage had yet to be drawn. At that time, also, the now cut-and-dried distinction between first-order and second-order language was not at all clear — a circumstance that might well have occasioned Zermelo's own intuition that something like his definiteness requirement was called for in order to rule out uses of separating formulae $\varphi(x)$ that otherwise might have been (at least, mathematically) obscure in content. A final drawback of the rule form just suggested, with its embedded turnstiles, is that one would no longer have an effective method of telling, when given an arbitrary sentence, whether it counts as an instance of the Axiom Scheme of Separation. Thus the ultimate set of axioms for set theory would be non-recursive. This complaint, however, is anachronistic, since the tradition had not yet produced that requirement of recursiveness upon axiomatic bases for formalized theories. The proponent of the rule form just suggested could also niggle in reply that proposed applications of that rule *are* actually effectively checkable for correctness. For, it could be argued, one would be able to check, effectively, whether the purported subproof of the definiteness of φ is indeed a proof. This niggle-room, of course, would be purchased only at the cost of conflating object language and metalanguage in the very formulation of the rule.

The post-Zermelo developments that have been mentioned — the object language/metalanguage distinction; the first-order versus second-order distinction; and the requirement of recursiveness for sets of axioms — enable one to see how, over the two decades or so after Zermelo first ruminated about definiteness, that requirement on instances of Separation simply lapsed. The underlying logic, being classical, rendered it toothless. Once set theory was fully first-orderized, the motivation for the requirement ebbed away; and once the object-meta distinction was firmly in place, axioms could not have turnstiles or anything redolent of the metalevel embedded within them.

Let us return, then, to the more measured and *constructivist* formulation of the rule for Separation, concerning which it is *not* assumed that one will have φ's definiteness-by-default thanks to one's logic being classical.

Because it is so much more easily parsed, only the rule form for Separation will be used in what follows. But first we provide another historical interlude.

1.3. Second Historical Interlude: Brouwer

The point has been reached in this closer than usual attention to Zermelo's own remarks about Separation, and his requirement that the separating formula be *definit*, where it is timely to address a question most probably already arising in the reader's mind (especially after reading the Fraenkel quote above). This is: *To what extent might this requirement of Zermelo have been inspired, or motivated, at least in part by any influence that Brouwer's early intuitionistic thinking might have had on him?*

It is tempting to think, and plausible (in light of the foregoing reflections on the culprit abstraction principle at work in the Goodman–Myhill result) that Brouwer's criticism of the usual (classical, default) formulation of Separation would almost certainly have been that it is premised on Bivalence — on the dogma that every mathematical question has a definite answer, *Yes* or *No*.[14] Yet Zermelo *seems* to have been unwilling to concur with this dogma; he requires that one actually furnish *proof* of the requisite *Definitheit* before Separation can be applied. Moreover, Brouwer himself *distinguished*, at least as separate misgivings, his reservations about the *Komprehensionsaxiom*, from those he had about Bivalence. In [Brouwer, 1919, pp. 203–204] he wrote the following passage, which is extraordinarily interesting in the context of the present discussion, but of which, to the best of the author's knowledge, no translation into English has been published.[15] The following is the author's translation.

---

[14] Note that 1908 saw the publication not only of Zermelo's axiomatization of set theory, but also of (the original Dutch version of) Brouwer's clarion rejection of *tertium exclusum*, in his paper 'The unreliability of the logical principles'. (See [1908].) For Brouwer,

> ... the question of the validity of the *principium tertii exclusi* is equivalent to the question *whether unsolvable mathematical problems can exist.*

Brouwer thought there could be such problems; but one might assume that Zermelo, under the powerful influence of Hilbert at Göttingen, would more likely have been a 'Hilbertian optimist', believing in the solvability of all mathematical problems, and in one's consequent entitlement to the Law of Excluded Middle.

[15] The original German reads as follows:

> Seit 1907 habe ich in mehreren Schriften [fn] die beiden folgenden Thesen verteidigt: I. dass das *Komprehensionsaxiom*, auf Grund dessen alle Dinge, welche eine bestimmte Eigenschaft besitzen, zu einer Menge vereinigt werden (auch in der ihm später von ZERMELO gegebenen beschränkteren Form [fn]) zur Begründung der Mengenlehre unzulässig bzw. unbrauchbar sei und nur in einer *konstruktiven* Mengendefinition eine zuverlässige Basis der Mathematik gefunden werden könne;
> II. dass das von HILBERT 1900 formulierte *Axiom von der Lösbarkeit jedes Problems* [fn] mit dem *logischen Satz vom ausgeschlossenen Dritten* äquivalent sei, mithin, weil für das genannte Axiom kein zureichende Grund vorliege and die Logik auf der Mathematik beruhe under nicht umgekehrt, der logische Satz vom ausgeschlossenen Dritten ein *unerlaubtes* mathematisches Beweismittel sei, dem kein andrer als ein scholastischer und heuristischer Wert zugesprochen werden könne, so dass Theoreme, bei deren Beweis seine Anwendung nicht umgangen werden kann, jeden mathematischen Inhalt entbehren.

> Since 1907 in several writings [fn] I have defended the following two theses:
>
> I. that the *Axiom of Comprehension*, on whose basis all things that possess a certain property can be united to form a set (also in the more constrained form given to that axiom later by Zermelo) is inadmissible or rather useless, and an admissible basis for mathematics will be able to be found only in a *constructive* definition of set;
>
> II. that the *Axiom of the solubility of every problem* [fn] formulated in HILBERT 1900 is equivalent to the logical Law of Excluded Middle; therefore, because no sufficient ground is at hand for the named axiom, and logic depends on mathematics and not the other way round, the logical Law of Excluded Middle is an *impermissible* means of mathematical proof, to which nothing but a scholastic and heuristic value can be ascribed — so that theorems for whose proof its application cannot be avoided are deprived of any mathematical content.

One can glean from this some reason to conjecture that Brouwer was concerned to stress the year 1907 because it was the year *before* the publication [Zermelo, 1908] to which Brouwer refers in the second footnote indicated, which reads 'Vgl. Math. Ann. 65, S. 263.' Brouwer's reference to this particular page of Zermelo's paper is clearly intended to focus the reader's attention on Zermelo's AXIOM III of Separation — which Brouwer characterizes as a '[beschränktere] Form' of the *Komprehensionsaxiom* (a restricted form of the Comprehension axiom).[16] Brouwer then proceeded (*loc. cit.*, pp. 204–205) to offer his own, constructive, definition of what a set (*Menge*) might be, which is based on laws (*Gesetze*) governing the generation of sequences of natural numbers.[17] All one needs to know, for its relevance to the present discussion, is that it is a conception of set devoid of any dependency on prior sets to which any operations of 'separation' might then be applied.

---

[16] This observation is all the more plausible in light of the fact that seven years earlier, in his inaugural address [1912] at the University of Amsterdam, Brouwer had mentioned a certain modified form of Separation (which he then called 'the axiom of inclusion'):

> If for all elements of a set it is decided whether a certain property is valid for them or not, then the set contains a subset containing nothing but those elements for which the property does hold. (At p. 90 in the English translation [Brouwer, 1975] in Brouwer's *Collected Works.*)

To the words here quoted Brouwer appended a footnote that reads 'Compare Zermelo, *Math. Annalen*, vol. 65, p. 263.'

[17] It is unnecessary to quote Brouwer's definition in full. In this 1919 paper he was essentially reprising (with some clarificatory elaboration) the definition of *Menge* that he had given earlier in one of the papers to which he referred in the second footnote indicated. This earlier definition will be found in the second paragraph of [Brouwer, 1918].

The record thus far, then, shows Brouwer responding with implied criticism, in 1919, to Zermelo's Axiom of Separation of 1908, while taking the opportunity to stress 1907 as the year in which he initiated his (Brouwer's) own departure from the classical conception of set. Whether that departure, that early on, might have been in reaction to anything that Brouwer had learned about Zermelo's developing thoughts about axiomatic set theory is a question for whose answer there is little definitive evidence.

The historical record appears to contain no correspondence between Brouwer and Zermelo. There is one very cordial letter from Brouwer to Hilbert, dated October 15th, 1909, in which Brouwer reminds Hilbert of a promise of his that he would send Brouwer a copy of Zermelo's 1908 paper:

> Allow me to remind you of your promise, to have reprints of yours (on the Dirichlet principle, integral equations and the Obituary of Minkowski) and of Mr. Zermelo (Foundations of set theory) and Koebe (uniformisation) sent to me? I express in advance many thanks. [van Dalen, 2013, p. 146]

There is also the following comment by van Dalen in his introductory note to Zermelo's 1937 piece 'Relativism in set theory and the so-called Skolem theorem':[18]

> It should be noted that Zermelo had another *bête noire*: intuitionism. Zermelo's references to this topic are not systematic enough to allow us to reach definite conclusions on his views. It seems that on the whole Zermelo did not wish to see intuitionism as a serious foundational program. ... On the whole, Zermelo seems to have been inadequately informed about Brouwer's doctrines; he accuses Brouwer, for example, of banning the infinite from mathematics. This may explain his harsh view of intuitionism. [Ebbinghaus *et al.*, 2010, pp. 602–603]

So it would appear to be unlikely that Brouwer's interest in [Zermelo, 1908] would have been by way of follow-up to any earlier exchange Brouwer might have had with Zermelo about *Definitheit*, with a resulting interest in how Zermelo might have responded to general Brouwerian concerns about decidability, *etc.*

In Brouwer's own selected correspondence, there is but one trace of awareness, on the part of one of Brouwer's correspondents, of the possible relevance of Bivalence to the *Definitheit* of Zermelo's separating formulae. A letter from Heinrich Scholz, dated November 16th, 1927, which appears to be a long summary of Brouwer's foundational views, contains the following interesting comments:[19]

---

[18] I am grateful to Mark van Atten for these two references.

[19] This is a faithful quote, without corrections of any of its strange punctuation marks. It is also worth noting that Ebbinghaus, in his Introduction 'Ernst Zermelo: A glance at his life and work' to [Ebbinghaus *et al.*, 2010], describes Scholz as 'Zermelo's best

> THE THREE MAIN FAILURES OF FORMAL LOGIC
> (1) The misuse of Tertium non Datur:
> consisting of
> a) the illegitimate application to arbitrary properties of a given individual,
> ...
> (2) the abuse of the notion of class, resp. property.
> consisting of the use of the above for the creation of non-constructible sets, and in particular totally unrestricted. or, as this unrestricted use has led to logical 'catastrophes, *under the determined conditions of the sharpened axiom of separation of Zermelo.* [van Dalen, 2013, p. 319, emphasis added]

Clearly, this connection descried by Scholz comes far too late to be any evidence for influence that Brouwer might have exercised on Zermelo in the latter's 1908 requirement of the *Definitheit* of separating formulae.

1.4. The 'Orderism' Issue

Note that the pursuit of an answer to the question about how best to understand Zermelo's notion of *Definitheit* is orthogonal to the discussion of *Definitheit* in [Shapiro, 1991, pp. 181–183]. Shapiro concerns himself only with the criticisms of Russell and Fraenkel (of Zermelo's notion of *Definitheit*), and suggested explications of it from Weyl and Skolem. The latter engage one only in the question whether the quantifiers within a separating formula may be permitted to be second-order. This of course is in keeping with the major theme of Shapiro's book; but it does not speak to specifically *intuitionistic* concerns, since these are invariant across answers to the 'orderism' question.

Zermelo, by the way, was unrelenting in his opposition to Skolem's recommended restriction to first-order quantifiers. As if to underscore this, in his introductory note to [Zermelo, 1929], Ebbinghaus [2010, p. 355] offers the following English rendition of an important passage:

(1) The set of definite propositions contains the fundamental relations and is closed under the operations of negation, conjunction, disjunction, first-order quantification, *and second-order quantification.* [Emphasis added.]

It is worth noting that the material italicized here in Ebbinghaus's gloss does not correspond literally to any part of Zermelo's original German for his (1). (Ebbinghaus's subsequent translation of the whole paper, however, is faithful to the original text.) Ebbinghaus's inclusion of 'and second-order quantification' is exegetically justified by Zermelo's earlier clause (4) inhis definition of how

friend and scientific partner' (p. 32). So one is led to wonder whether Scholz might have communicated any of his own understanding of these aspects of intuitionism to Zermelo.

Definitheit is transmitted to compound sentences from their constituents. That particular clause, in the original, reads as follows (where '$D\ldots$' stands for '$\ldots$ ist definit'):

> Gilt $DF(f)$ für alle definiten Funktoren $f = f(x, y, z, \ldots)$, so gelten auch $D\bigcap_f F(f)$ und $D\bigcup^f F(f)$.

One can see that Zermelo is insisting here on having recourse to second-order quantifiers (over *Funktoren*) in his separating formulas.

Zermelo could not have been aware (for this was the year of Gödel's *Doktorarbeit*) that, while first-order deducibility captures first-order logical consequence, second-order logical consequence, by contrast, cannot be captured by any system of finitary proof. For Zermelo to insist on admitting second-order separating formulae, *and cleaving to his 1908 notion* of *Definitheit*, is to open up the possibility of having (from the classical realist's point of view) certain subsets of a given set separable from it by means of formulae that fail to be *definit* by the lights of Zermelo's own requirement in his comments on his Axiom of Separation. *This* is the possibility on which the constructivist would have to seize, in order to ensure the non-triviality of Zermelo's requirement on applications of Separation.

1.5. Back to the Logical Analysis of the Role of the Goodman–Myhill Abstraction Principle

Let us return now to the construal bruited earlier, of the culprit abstraction principle really being an easy consequence of a substitution instance of the Axiom Scheme of Separation *in its usual contemporary formulation.* If, instead, the amended formulation of Separation proposed above (§1.2) were to be insisted upon, one would be invoking as supposedly licit (in the context of the regimentation here of the Goodman–Myhill proof, and the actual application it makes of Separation) a step of the form

$$\dfrac{\exists!\{0\} \qquad \begin{array}{c} \overline{a \in \{0\}}^{(i)} \\ \vdots \\ (a \in \{0\} \wedge \psi) \vee \neg(a \in \{0\} \wedge \psi) \end{array}}{\exists!\{x \in t|\psi\}}{\scriptstyle (i)},$$

where the subformula ψ is a *sentence*, *i.e.*, has no free variables, and the parameter a can be chosen so as not to occur in ψ. And the right-hand subproof here, given its assumption $a \in \{0\}$ that is made for subsequent discharge, can be constructed in only two possible ways:

$$\dfrac{\dfrac{a\in\{0\} \qquad \begin{array}{c} a\in\{0\} \\ \vdots \\ \psi \end{array}}{a\in\{0\}\wedge\psi}}{(a\in\{0\}\wedge\psi)\vee\neg(a\in\{0\}\wedge\psi)} \qquad \text{or} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\overline{a\in\{0\}\wedge\psi}^{\,(1)}}{\psi} \qquad \begin{array}{c} a\in\{0\} \\ \vdots \\ \neg\psi \end{array}}{\bot}}{\neg(a\in\{0\}\wedge\psi)}{\scriptstyle (1)}}{(a\in\{0\}\wedge\psi)\vee\neg(a\in\{0\}\wedge\psi)}\ .$$

In the two subproofs indicated by vertically descending dots, the parameter a does not occur in ψ, nor in any assumptions, other than $a \in \{0\}$, on which those subproofs' conclusions ψ and $\neg\psi$ respectively depend. (Among those a-free assumptions, of course, there can be axioms of one's set theory.) One can therefore substitute 0 for a. Since $0 \in \{0\}$ is a *theorem* of the set theory, this means that one will achieve, respectively, a proof of ψ, or a proof of $\neg\psi$, as a theorem of the set theory.

It follows that for the amended form of Separation to be invoked in the Goodman–Myhill proof, one has to be able either to *prove* ψ, or to *refute* it. For the constructivist, that means: prove $\psi \vee \neg\psi$.

It is now absolutely clear that, for one who takes Zermelo's insistence on *Definitheit* seriously, the Goodman–Myhill 'proof' of $\psi \vee \neg\psi$ is utterly circular, as an argument designed to persuade *constructive* Zermelo set-theorists that they are committed to accepting Excluded Middle across the board. The sentence ψ has to be decidable in order for the amended form of Separation to be able to involve it in the way that is required by the Goodman–Myhill proof strategy. So the conclusion of their proof — $\psi \vee \neg\psi$ — comes as no surprise at all. Indeed, it is innocuous, given that this has to be the case in order for Separation (in its amended form) to be applicable.

This diagnosis also has the following signal advantage. It fits well with a 'straightforward' constructivizing approach to classical mathematical theories. On this approach one carefully chooses constructively acceptable forms of the mathematical axioms in question. (These forms are equivalent for the classicist, who therefore does not care which particular form is taken.) One then seeks to 'constructivize' simply by hewing to constructive logical reasoning from the chosen axioms, rather than indulging in non-constructive classical reasoning.

Had Zermelo's own insistence on *Definitheit* for the separating formula used in applications of his Axiom Scheme of Separation been respected — by adopting the appropriately modified form(s) proposed above — then the 'straightforward' constructivizer of classical set theory would have remained unmoved by the Goodman–Myhill 'result', as it has been regimented in this study.

2. SEPARATION IN INTUITIONISTIC AND/OR CONSTRUCTIVE SET THEORIES AFTER BISHOP

The foregoing considerations lead one to question whether some well-known axiomatizations of intuitionistic or constructivist set theory are actually acceptable to the intuitionist or constructivist respectively (for both of whom, of course, Excluded Middle is anathema).[20] It is worth looking at how various authors in pursuit of an adequate set-theoretic foundation for the constructivist mathematics of [Bishop, 1967] have treated Separation. They will be considered in roughly chronological order.

2.1. Myhill

It is somewhat ironic to have to include [Myhill, 1975] here. In this paper, appearing three years before [Goodman and Myhill, 1978],[21] the Axiom Scheme of Separation receives the following statement (pp. 350–351). Myhill deploys a formalism in which numbers, functions, and sets form pairwise disjoint sorts. The number-sort is called $\boldsymbol{Z}$.

> Now comes *predicative* separation (sometimes called Δ_0-separation), i.e.,
> D5 $\qquad (\exists X)(\forall x)(x \in X \Leftrightarrow x \in A \wedge \phi^a_x)$
> where every bound variable of ϕ is restricted to a set. Formally a *restricted formula* is defined as follows: atomic formulas are restricted, propositional combinations of restricted formulas are restricted, and if ϕ is restricted and t is a parameter or $\boldsymbol{Z}$, $(\forall x \in \mathrm{t})\phi^a_x$ and $(\exists x \in \mathrm{t})\phi^a_x$ are restricted.

Since $\boldsymbol{Z}$ is a set, Myhill is hereby allowing for separating formulae that are undecidable, even though they are 'predicative' in his explicated sense. And that is enough to make his 1975 system objectionable on the grounds that will be articulated in §3.

2.2. Friedman

For the inclusion of full-blown Separation in a set of axioms purporting to capture intuitionistic set theory see [Friedman, 1973]. At p. 216 he writes

> Let ... A vary through all formulae in the language of ZF in which x, y, z do not occur. ... The axioms of ZF are ...
> 4. *Separation.* $(\exists x)(\forall y)(y \in x \leftrightarrow (y \in a \ \&\ A^b_y))$. ...

All the axioms of ZF that Friedman lists (with the exception of Extensionality) are retained as axioms for the intuitionistic set theory into which Friedman furnishes a syntactic transformation of ZF. In particular, this very form of Separation just quoted is an axiom (scheme) for his intuitionistic set theory.

---

[20] One can set aside, as orthogonal to the current line of inquiry, the fact that Brouwerian intuitionists prove certain theorems that contradict those of classical mathematics, while Bishop-style constructivists only cut back on what is classically provable, by eschewing Excluded Middle while still using only such axioms as the classical mathematician would accept.

[21] Laura Crosilla has pointed out (personal correspondence) that [Myhill, 1975] includes in the list of references an item 'to appear', that must have been what later became [Goodman and Myhill, 1978]. It bore the title 'The axiom of choice and the law of excluded middle'.

Note that there is no attempt to impose any requirement of *Definitheit* on the separating formula A.

Friedman provided two translation mappings $*$ and $-$ on formulae, and established that for every theorem φ of classical ZF (*not*: ZFC!), the translated correlate $*(-(\varphi))$ can be derived *intuitionistically* from the axioms ZF *minus* Extensionality. The mapping $-$ is a double-negation translation mapping of the familiar kind; while the mapping $*$ involves replacing each atomic subformula of the form $a \in b$ by an extraordinarily complex and ingeniously devised formula $a \in^* b$ with the same free variables.

Friedman's result (that 'classical set theory' is consistent relative to 'a set theory with intuitionistic logic') is clearly to the effect that classical *Zermelo–Fraenkel* set theory is consistent relative to a set theory closed under only intuitionistic logic, to be sure, but equipped with full-blown Separation. His method of proof (of relative consistency) cannot, in light of the Goodman–Myhill result, be generalized so as to show (as a non-trivial result) that classical ZFC is consistent relative to the (supposedly properly contained) intuitionistic closure of the ZFC axioms.

In the later paper [Friedman, 1977] greater caution is exercised. Friedman formulates axioms for a system **B** of constructive set theory, designed explicitly as a *formal system* to capture the (informal) *framework* within which the treatment of constructive analysis is conducted in [Bishop, 1967]. The formal system is 'Kroneckerian', in that it takes the natural numbers as given (as 'Urelemente'), and not needing to appear in the guise of set surrogates before being granted existence. Indeed, natural numbers *are not sets.* One of Friedman's 'Ontological Axioms' says that everything is a set or a natural number, but not both.[22] The system **B** also furnishes the set of all natural numbers, as a completed totality: its Axiom of Infinity says that the set of all natural numbers exists. (This is in the methodological spirit of Friedman's pioneering work in Reverse Mathematics.) Without any discussion of the reasons why, Friedman (like Myhill before him) here restricts the Axiom Scheme of Separation by requiring that the separating formulae be Δ_0. He also limits himself to Dependent Choice (not full Choice). Following Myhill, he eschews the Power Set Axiom, and uses instead the Exponentiation Axiom, remarking (at p. 6) that

> There is no essential use of the power set axiom in [Bishop's book]: what is used instead is the existence of the set of all total functions from one set into a second (the axiom of exponentiation).

2.3. Aczel

In setting up his system CZF of constructive set theory, Aczel [1978, p. 57] stated an axiom scheme of what he called Restricted Separation:

[22] Here, the formal system is very faithful to Bishop's framework, in which the naturals are treated as *sui generis*.

For restricted $\phi(\mathrm{x})$

$$\exists \mathrm{z}[(\forall \mathrm{y} \varepsilon \mathrm{z})(\mathrm{y} \varepsilon \mathrm{x} \wedge \phi(\mathrm{y})) \wedge (\forall \mathrm{y} \varepsilon \mathrm{x})(\phi(\mathrm{y}) \rightarrow \mathrm{y} \varepsilon \mathrm{z})].$$

Here, a restricted formula is one in which all quantifications are of the restricted forms $(\forall \mathrm{x} \varepsilon \mathrm{y})\psi(\mathrm{x})$ or $(\exists \mathrm{x} \varepsilon \mathrm{y})\psi(\mathrm{x})$. Clearly such restriction does not eliminate the possibility that the separating formula ϕ will be effectively undecidable.

2.4. Feferman

In explaining '[g]eneral features of the constructivist conception', Feferman [1979, p. 163] writes:

> <u>Sets</u> are only given by <u>defining</u> <u>properties</u>, for which we are supposed to know and understand their condition for membership. ... if A is any set and $\phi(\mathrm{x})$ is a well-understood property of members of A then $\mathrm{B}=\{\mathrm{x}\in\mathrm{A}|\phi(\mathrm{x})\}$ is a set, with $\mathrm{x}\in\mathrm{B} \leftrightarrow \mathrm{x}\in\mathrm{A} \wedge \phi(\mathrm{x})$.

Feferman does not say, however, what it is for a property to be 'well-understood'.

2.5. Beeson

In this regard, compare [Beeson, 1979, pp. 6–8]:

> ... we describe the principal intuitionistic set theories, which have been invented and studied by Friedman and Myhill. First we describe Friedman's systems. ...
>
> Let ZF^- be Zermelo–Fraenkel set theory, with intuitionistic logic, and with the foundation axiom expressed as transfinite induction on $\in$, instead of the usual way. (The usual foundation axiom implies the law of the excluded Middle, see [[Myhill, 1975]].) We cannot add the axiom of choice AC without getting the law of the excluded middle, but we can add (some forms of) dependent choice. The strongest set theory we consider is thus ZF^-+RDC (relativized dependent choice). (Introduced in [Friedman, 1973].) ... We now list the axioms we will be considering; we give them first in the form suitable when extensionality is present. ...
>
> M. (separation) $\exists x \forall y (y \in x \leftrightarrow (y \in a \wedge Q)$, where x is not free in Q.

Note that Beeson imposes no significant restrictions on Q (to ensure either constructivity or predicativity) in this scheme for Separation. So the extension of any system containing (M) by adopting Choice will yield Excluded Middle.

The slightly later paper [Beeson, 1981, p. 160], echoes the quote from Feferman, with 'well-defined' in place of 'well-understood':

> Subset existence. If W is a set and ϕ is any well-defined property such that $x =_W y$ & $\phi(x) \rightarrow \phi(y)$, then S = {x∈W:$\phi$(x)} exists.

Beeson does not say, however, what it is for a property to be 'well-defined'.

In the further paper [Beeson, 1982, pp. 13–14] the reader is introduced to the difference between *presets* and *sets*:

> In the current terminology of the new constructive mathematics, a set X is considered defined when we have
>
> (i) said what must be done to construct a member of X.
> (ii) said what must be done to prove two members of X are equal.
> (iii) proved that equality on X as defined in (ii) is an equivalence relation.
>
> If only (i) has been done, we say that X is a *preset*.

Beeson continues, at pp. 39–40:

> Certain principles of set existence are rather widely accepted. Among them are: ...
> *Separation:* If A is a preset and ϕ is a predicate on A then {x∈A:ϕ(x)} is a preset.

There is, conspicuously, no real restriction here on what would count as a 'predicate'. Presumably what Beeson means here is any well-formed formula, with just x free, in the language of set theory.

2.6. Greenleaf

Interestingly, [Greenleaf, 1981, pp. 218–219] comes closest to echoing Zermelo's original *caveat* about separating formulae:

> Let X be a set, and A a predicate on X. Corresponding to A we have the subset S={x|A(x)}. To construct an element of S it is necessary first to construct an element x of X, **and then to prove A(x)**. [Boldface emphases added.]

It is clear that the extension of the predicate A has to be included within the set X (otherwise, presumably, S as Greenleaf defines it would not be a subset of X). Other than that, he gives no further guidance as to what, exactly, the predicate A can be.

2.7. McCarty

McCarty [1986], at p. 158, writes

> IZF is formulated in a single-sorted first-order language with $\in$ and $=$ as its only primitive nonlogical predicates. The axioms of IZF are all

> instances of ...
> [SEP] $\forall x \exists y \forall z(z \in y \leftrightarrow (z \in x \wedge \phi))$ ...

and presumably in this axiom scheme ϕ may be any sentence, or formula with z as its only free variable. McCarty continues:

> In classical logic, the axioms of IZF are equivalent to traditional Zermelo–Fraenkel. In intuitionistic logic, this equivalence fails; IZF does not derive the general law of excluded third.

McCarty's full list of axioms contains, besides Separation, those of *Extensionality*, Pairing, Union, Power Set, Infinity, Collection, and Induction on $\in$. It therefore follows, by the Goodman–Myhill result, that the intuitionistic theorist cannot contemplate extending IZF, on McCarty's presentation, by adding the Axiom of Choice (as McCarty himself was no doubt well aware). In IZF, Separation is simply too *carte blanche*. It has the instance

$$\forall z \exists y \forall x(x \in y \leftrightarrow (x \in z \wedge \psi)),$$

that is crucially at work — in combination with Choice — in precipitating Excluded Middle via the Goodman–Myhill argument. The reader will recall that all it takes is one step of $\forall$E on this instance of Separation, to get

$$\exists y \forall x(x \in y \leftrightarrow (x \in \{0\} \wedge \psi)),$$

and then, because of Choice, it would be off to the classical races. Naturally, McCarty refrains from adopting Choice, for this very reason. Indeed this is the consensual response among constructive set theorists: retain even unrestricted Separation, but give up Choice, so as to avoid commitment to Excluded Middle.

2.8. Troelstra and van Dalen

In Volume II of their monumental study of constructive mathematics — [Troelstra and van Dalen, 1988] — the authors frame Restricted Separation (for CZF) as follows:

R-Sep $\qquad \forall x \exists z \forall y(y \in z \leftrightarrow y \in z \wedge A[y]) \qquad$ (A restricted).

Of course the final occurrence of z should be an occurrence of x. The by now familiar observation applies in their case: provided they have the set of natural numbers in their ontology, they will have undecidable separating formulae.

2.9. Aczel and Rathjen

In the draft book [Aczel and Rathjen, 2010], the authors state various axioms and axiom schemes constituting the axiomatic bases of various systems of constructive set theory. Their preferred system is called **CZF** (for Constructive Zermelo–Fraenkel) and it has the following axiomatization. (On p. 25 they had explained 'A formula is bounded if all its quantifiers are bounded; i.e. occur only in one of the forms $\exists x \in y$ or $\forall x \in y$.')

Extensionality	$\forall x \forall y(\forall z(z \in x \leftrightarrow z \in y) \rightarrow x = y$
Pairing	$\forall x \forall y \exists z \forall w(w \in z \leftrightarrow (w = x \vee w = y))$
Union	$\forall x \exists y \forall z(z \in y \leftrightarrow \exists w(z \in w \wedge w \in x)$
Strong Infinity	$\exists z(Ind(z) \wedge \forall y(Ind(y) \rightarrow \forall w(w \in z \rightarrow w \in y)))$
Bounded Separation	$\forall z \exists y \forall x(x \in y \leftrightarrow (x \in z \wedge \varphi(x)))$ for bounded $\varphi(x)$ not containing y free
Strong Collection	$\forall z(\forall x(x \in z \rightarrow \exists y \varphi(x,y)) \rightarrow \exists w[\forall x(x \in z \rightarrow \exists y(y \in w \wedge \varphi(x,y)))$ $\wedge \forall y(y \in w \rightarrow \exists x(x \in z \wedge \varphi(x,y)))])$ for all formulae $\varphi(x,y)$ not containing w free
Set Induction	$\forall w(\forall x(x \in w \rightarrow \varphi(x)) \rightarrow \varphi(w)) \rightarrow \forall z \varphi(z)$ for all formulae $\varphi(w)$
Subset Collection	$\forall z \forall v \exists w \forall u[\forall x(x \in z \rightarrow \exists y(y \in v \wedge \psi(x,y,u))) \rightarrow$ $\exists d(d \in w \wedge \forall x(x \in z \rightarrow \exists y(y \in d \wedge \psi(x,y,u)))$ $\wedge \forall y(y \in d \rightarrow \exists x(x \in z \wedge \psi(x,y,u))))]$ for every formula $\psi(x,y,u)$ not containing w free.

Here, $Ind(w) \equiv_{df} \exists y(y \in w \wedge \forall z \neg z \in y) \wedge \forall x(x \in w \rightarrow \exists y(y \in w \wedge Succ(x,y)))$, where $Succ(x,y) \equiv_{df} \forall z(z \in y \leftrightarrow (z \in x \vee z = x))$.

Once again, one sees Separation, albeit in Bounded form, in too strong a formulation — one that admits the culprit instance used in the Goodman–Myhill proof of Excluded Middle from Choice. That is why Aczel and Rathjen omit Choice.

2.10. Upshot of the Survey of Constructive Set Theorists

The upshot of all these constructivist attempts to furnish a *constructive* set theory seems to be an apparent unwillingness to subject Separation to the degree of skeptical constructivist critique that this author holds would be warranted. There appears to have been no attempt to *constructivize the notion of Definitheit* that Zermelo so intriguingly saw fit to impose on the separating formulae involved in his Axiom Scheme of Separation.

The overall predicament is clear: there is, essentially, a constructive proof of the Law of Excluded Middle from the combination of *two* salient premises: a certain form of the Axiom of Choice; and a particular instance of Separation. The Law of Excluded Middle is to be avoided at all costs; so the constructivist must give up one or other of those two premises. But which one?

This is a simple problem of belief-revision — simple in form, at least, if not in solution. When the combination of certain earlier beliefs yields a patently unacceptable result, one may need to import some pragmatic considerations, such as relative entrenchment, or conformity with important guiding philosophical conceptions. As already seen, there is actually a reasonably persuasive argument *in favor of* the Axiom of Choice, based on the constructivist's understanding of the logical operators involved. This point is made by both Bishop and Dummett. And on closer inspection, it appears that uncritically following the classical set theorist's preferred choice of logical form for Separation is where the real problem lies. It may well be that a more fully enlightened constructivist outlook — prompted by more careful logical reflection on the Goodman–Myhill result — would raise qualms about accepting, as a first principle,

$$\exists y \forall x(x \in y \leftrightarrow (x \in \{0\} \wedge \psi)),$$

for *absolutely any* sentence ψ (in the language of first-order set theory) that the scheme instantiater cares to employ. Even if this is to venture beyond even Brouwerian strictures, it may be time to ask what grasp the constructivist can really claim to have of the set thus claimed to exist — that is, of *what its members are* — when for ψ there is chosen (the set-theoretical expression of) Goldbach's Conjecture; or the Twin-Prime Conjecture; or the Riemann Hypothesis; or the Continuum Hypothesis . . . .

2.11. Which to Prefer: Choice or Separation?

Throughout this section, 'Choice' is to be read as 'full Choice'. More restricted forms of Choice will be referred to by their more specific labels, such as AC^{ω}.

The dialectic that is emerging is one that could well divide constructivists. Their main representatives have not really tried to impose any significant degree of *Definitheit* on their separating formulae for applications of Separation.[23] And there is a quite firmly established constructivist tradition now of eschewing Choice as a *non*-constructive principle. The non-constructivity they object to manifests itself mainly in the context of such set theory as would be required for a constructive foundation for real analysis. There, the adoption of Choice rules out the possibility of having every real function be continuous — a possibility that the post-Bishop constructivist, while not necessarily being willing to actualize it (on behalf of the Brouwerian intuitionist), is not willing to rule out, either. The resulting reluctance to countenance Choice is intimately tied to the significant intellectual leap that has to be taken to get from the realm of finite sets to the full Cantorian realm — Hilbert's 'paradise' — that contains, in addition, (uncountably) infinite sets.

Since 1978, however, the reluctance to countenance Choice has been reinforced by the Goodman–Myhill result. It appears to draw on meager set-theoretical reserves in deducing Excluded Middle from Choice, to the supposed discredit of the latter. It *appears* not even to need the surroundings of infinite sets to give the constructivist pause about Choice. The rout of Choice, as a potential fixture on the constructivist scene, appears to be complete. It is worth reminding the reader that our proof-analysis has revealed that the form of choice actually called for in a fully detailed regimentation of Goodman and Myhill's proof is what we have called GM-Choice, which appears to be much weaker than Full Choice.

The argument for GM-Choice on the basis of the BHK readings of the logical operators involved acquires great plausibility in the imagined situation of a 'Kroneckerian constructivist', who is concerned to theorize only about the

[23] Friedman, with his system **B**, is a possible exception here; he requires the separating formula to be Δ_0.

hereditarily finite pure sets, and who refuses to acknowledge the existence of any 'completed' infinite set (such as ω, the set of all finite von Neumann ordinals). And, as already seen, close analysis of the Goodman–Myhill proof reveals that it is not Choice alone that visits Excluded Middle upon the constructive set theorist. It is also Separation.

The reader is invited to consider whether Separation is not a mere partner in crime, let alone a non-offending bystander. It might, rather, be the real culprit — perhaps most clearly when one considers the most appropriate setting to make uncluttered comparison. By 'most appropriate setting' is meant: that constructive theory of sets that could provide a foundation on which to erect the *Arithmetik* of the Kroneckerian constructivist, for whom there are no completed infinite totalities. The reader is invited to explore the issue at hand — Choice *versus* Separation — while occupying the shoes, and the vantage point, of the Kroneckerian constructivist. As Kronecker famously said, God gave us the integers; humankind did all the rest. Let us unwrap the divine gift as the finite von Neumann ordinals (but refrain from thinking there is a set of all of them!);[24] and let us investigate further just how the constructive set theorist might set about doing as much of 'the rest' as is constructively possible.

In making a case, however, for a restricted form of Separation even when 'living within' just the Kroneckerian set-theoretic universe of hereditarily finite pure sets, it might be instructive to take the liberty of advancing metamathematical considerations that are available only to the foundationalist who occupies a more expansive universe — one that includes, among other things not countenanced by the Kroneckerian, the completed set ω of all finite ordinals. This does not weaken the constructivist's case; it should only strengthen it, in the eyes of those who decline to lead intellectual lives of Kroneckerian austerity.

The case to be argued is that Separation needs to be reined in — significantly. This is necessary in order to avoid having Excluded Middle thrust upon us even in the constructively comfortable climes of the hereditarily finite pure sets. If Separation has to be reined in *there*, then most surely the reins need to be kept at least that tight when one ventures to constructively more precarious heights, upon postulating the existence of ω. Once it is accepted that Separation may be used only under continuing constraint, one can begin to inquire what sorts of sets one might be able to recognize as existing upon accepting the existence of ω.

3. THE ABIDING ISSUE OF THE DECIDABILITY OF SEPARATING FORMULAE, FOR THE CONSTRUCTIVIST

The 'orderism issue' is the one joined by Skolem against Zermelo. Skolem insisted that the theory of sets — including, of course, and very importantly,

[24] The divine gift could also be unwrapped as *Urelemente*, taken severally, and *not* being identified as finite von Neumann ordinals, let alone as themselves forming a set.

the Axiom Scheme of Separation:[25]

$$\forall x \exists y \forall z (z \in y \leftrightarrow (z \in x \wedge \psi)), \text{ where } y \text{ is not free in } \psi$$

— should be given in a first-order language. In this formulation, ψ is a schematic placeholder for first-order formulae in the language of set theory. It may or may not contain z free. That is, either ψ is a *unary formula* $\psi(x)$, or it is a *sentence.* So the axiom scheme has infinitely many instances, but they are all at first order. Zermelo insisted, by contrast, that the theory should be formulated at second order, thereby allowing the Axiom of Separation to take the form of a single quantified sentence:

$$\forall \Psi \forall x \exists y \forall z (z \in y \leftrightarrow (z \in x \wedge \Psi)),$$

where $\forall \Psi(\ldots \Psi \ldots)$ is a second-order quantifier.

The debate over the orderism issue, joined by Skolem and Fraenkel against Zermelo, preceded three important formal discoveries — whose proofs, as already intimated, demand more than merely Kroneckerian resources. First, there was Gödel's completeness theorem (1929) for first-order logic. Second, there was Gödel's first incompleteness theorem for consistent and sufficiently strong first-order theories of arithmetic, along with his second incompleteness theorem to the effect that no such theory can establish its own consistency. Clearly,

(†) second-order Peano Arithmetic has the standard model $\mathbb{N}$ as its only model.[26]

The first incompleteness theorem had, as a corollary,

(i) there are non-standard first-order models for first-order Peano Arithmetic.

---

[25] What is said here about Separation applies also, of course, to Replacement — which implies Separation, even though Separation is usually stated separately when giving the axioms of ZF.

[26] The axioms for 0 and successor, along with the recursion axioms for addition and multiplication, ensure the correct extensions for those operations on the standard natural numbers. Moreover, the second-order axiom of mathematical induction:

$$\forall \Phi((\Phi(0) \wedge \forall x(\Phi(x) \rightarrow \Phi(s(x)))) \rightarrow \forall x \Phi(x))$$

ensures that all members of the domain are standard natural numbers. To see this, take for $\Phi(x)$ the instance 'x is at most finitely many steps of successor away from 0'.

Moreover, in light of (†), the first incompleteness theorem had as a further corollary

(ii) second-order logic has no sound *and complete* system of decidable proof.[27]

Third, there was the subsequent realization (not only that second-order logical truth is not axiomatizable, but also) that second-order logical consequence is not compact. Both of these second-order notions — logical truth, and compactness — are based on the full power-set semantics for second-order languages, rather than the Henkin semantics (for which, see [Henkin, 1950]). In the latter, proper subsets of the power set of the domain can serve as the range of variation for second-order quantifiers over properties.

Accordingly, by the late 1930s it was well known that first-order language maximized *deductive* power by perforce sacrificing *expressive* power; while second-order language maximized *expressive* power (yielding, for example, the categoricity of second-order Peano Arithmetic) by perforce sacrificing *deductive* power. This impossibility — of simultaneously maximizing both deductive and expressive power — turns out to be universal, affecting every possible language for mathematics. (See [Tennant, 2000].)

One must now inquire how these considerations position the constructivist who is wrestling with the problem of how best to formulate the Axiom (or Axiom Scheme) of Separation in set theory. To bring the matter into sharper focus, consider how constructivists might proceed if their intended *model* were the universe $\mathcal{HF}$ of *hereditarily finite* pure sets.

The *theory* HF would contain the *negation* of the usual Axiom of Infinity.[28] But all other axioms, and axiom schemes (if there are any of the latter) could remain as they are for ordinary ZFC (minus Infinity). Of such pairs of classically equivalent versions as there may be for certain axioms, the constructivist would of course choose the one that is constructively acceptable. Regularity implies Foundation constructively in ZFC, while the converse holds only classically. A final principle to mention as a valuable addition for *constructive* HF is that of $\in$-induction, which is a scheme, and which can be adopted in place of Regularity or Foundation:

$$\forall x(\forall w \in x(\varphi(w) \to \varphi(x)) \to \forall y \varphi(y).$$

[27] To see this, consider all logical truths of the form $\mathrm{PA}_2 \to \varphi$, where PA_2 is the conjunction of the (finitely many) axioms of second-order Peano Arithmetic, and φ is any first-order sentence in the language of arithmetic. In light of (†), any effective enumeration of all logical truths of this form would yield an effective enumeration of the first-order truths of Peano Arithmetic, contrary to the first incompleteness theorem.

[28] Friedman [2003, p. 26] uses the following Axiom of Finiteness:

$$\forall x(x \neq \emptyset \to (\exists y \in x)(\forall z \in x)(y \notin z)).$$

This principle is true in the intended model $\mathcal{HF}$ and turns out to be necessary in order for the mutual interpretations of first-order PA and HF to be inverses of each other.[29] *Modulo* the other axioms of HF, the principle of $\in$-induction is equivalent to the claim that every set has a transitive closure.

The proposal to consider HF rather than ZFC puts one in a good position to assess the Goodman–Myhill result from the point of view of the Kroneckerian constructivist, who was brought into the discussion in §2.11. Note, first, that the Axiom of Infinity plays no role in the Goodman–Myhill proof.[30] And note, second, that both Choice and Separation are true in $\mathcal{HF}$. That is, both of the culprit premises for the Goodman–Myhill proof of Excluded Middle are available to the Kroneckerian constructive set theorist.

Second-order HF has $\mathcal{HF}$ as its only model. Thus second-order HF logically implies (on the full power-set semantics) every sentence in the second-order language of set theory that is true in $\mathcal{HF}$. Among these are the *first-order* set-theoretic translations $\tau(\varphi)$ of all those sentences φ of the first-order language of arithmetic that are true in the standard model $\mathbb{N}$ of arithmetic. Among these first-order set-theoretic translations is that of $\mathrm{Con}_{\mathrm{PA}}$. Let us assume that it is true. Then the first-order set-theoretic sentence $\tau(\mathrm{Con}_{\mathrm{PA}})$ is a second-order logical consequence of second-order HF. Hence the formula

$$x = x \wedge \tau(\mathrm{Con}_{\mathrm{PA}})$$

would count as a *definit* separating formula in the Axiom of Separation, according to Zermelo's requirement that expressed *Definitheit* in the weaker terms of truth-on-instances being *determined* by the fundamental relations on the domain. But what about the version of *Definitheit* that was given in the stronger terms of *showing* or *proving* those instances? Given that second-order logic is incomplete, there is, *prima facie*, no guarantee that $\tau(\mathrm{Con}_{\mathrm{PA}})$ can be proved.

The predicament for the Zermeloan is even worse if one limits oneself to first-order theorizing, even though there is a complete proof system at first-order. For, given that first-order HF can be interpreted in first-order PA, we *know* that $\tau(\mathrm{Con}_{\mathrm{PA}})$ can be neither proved nor refuted in first-order HF.

This can be seen as follows. It is well known that (classical) HF and PA are *definitionally equivalent*, in a sense to be defined presently. Suppose this definitional equivalence is via the following translations of the language of one of those theories into the language of the other:

$$\tau : L_{\mathrm{PA}} \mapsto L_{\mathrm{HF}}; \qquad \tau' : L_{\mathrm{HF}} \mapsto L_{\mathrm{PA}}.$$

Here of course a language is being identified with the set of all its sentences. Now for the promised definition of definitional equivalence.[31]

---

[29] See [Kaye and Wong, 2007, p. 502]. This study shows just how much subtle work is required in order to confirm a result long thought to be part of 'folk lore'.

[30] Nor, for that matter, does the Axiom of Finitude!

[31] Note that this definition yields distinct intuitionistic and classical notions, depending on whether the turnstile is interpreted as intuitionistic or classical deducibility, respectively. A plausible conjecture is that HA (the intuitionistic closure of the first-order Peano

Theories T_1, T_2 (in languages L_1, L_2 respectively) are *definitionally equivalent* via translations $\tau_{12} : L_1 \mapsto L_2$ and $\tau_{21} : L_2 \mapsto L_1$ if and only if for all φ_1 in L_1, for all φ_2 in L_2, the following clauses hold:

(i) $T_1 \vdash \varphi_1 \Rightarrow T_2 \vdash \tau_{12}(\varphi_1)$;
(ii) $T_2 \vdash \varphi_2 \Rightarrow T_1 \vdash \tau_{21}(\varphi_2)$;
(iii) $\varphi_1 \in L_1 \;\Rightarrow\; [\tau_{21}(\tau_{12}(\varphi_1)) \dashv_{T_1} \vdash \varphi_1]$; and
(iv) $\varphi_2 \in L_2 \;\Rightarrow\; [\tau_{12}(\tau_{21}(\varphi_2)) \dashv_{T_2} \vdash \varphi_2]$

Let us pick up the thread now of explaining the independence of $\tau(\mathrm{Con}_{\mathrm{PA}})$ from HF. Consider the following instantiation of the general scheme in the foregoing definition of definitional equivalence:

$$\begin{array}{cccc} T_1 & T_2 & \tau_{12} & \tau_{21} \\ \downarrow & \downarrow & \downarrow & \downarrow \\ \mathrm{PA} & \mathrm{HF} & \tau & \tau' \end{array} .$$

By definitional equivalence (see clause (iii)) we have

$$\tau'(\tau(\mathrm{Con}_{\mathrm{PA}})) \dashv_{\mathrm{PA}} \vdash \mathrm{Con}_{\mathrm{PA}}.$$

Suppose for *reductio* that

$$\mathrm{HF} \vdash \tau(\mathrm{Con}_{\mathrm{PA}}).$$

Then by definitional equivalence (see clause (ii)) we have

$$\mathrm{PA} \vdash \tau'(\tau(\mathrm{Con}_{\mathrm{PA}})).$$

By Cut it now follows that

$$\mathrm{PA} \vdash \mathrm{Con}_{\mathrm{PA}},$$

contradicting Gödel's second incompleteness theorem. So by *reductio* we have

$$\mathrm{HF} \nvdash \tau(\mathrm{Con}_{\mathrm{PA}}).$$

Now suppose for *reductio* that

$$\mathrm{HF},\ \tau(\mathrm{Con}_{\mathrm{PA}}) \vdash \bot.$$

Then by $\neg$-I we have

$$\mathrm{HF} \vdash \neg\tau(\mathrm{Con}_{\mathrm{PA}}).$$

---

axioms) and the intuitionistic closure of appropriately constructive axioms for the first-order theory of hereditarily finite pure sets will turn out to be definitionally equivalent via appropriately specified translations.

It follows by the usual clause for negated sentences in the definition of τ that

$$\mathrm{HF} \vdash \tau(\neg\mathrm{Con}_{\mathrm{PA}}).$$

Hence by definitional equivalence (see clause (ii) again)

$$\mathrm{PA} \vdash \tau'(\tau(\neg\mathrm{Con}_{\mathrm{PA}})).$$

But by the definitional equivalence of PA and HF via the translations τ and τ' (see clause (iii)), we have

$$\mathrm{PA}\,,\ \tau'(\tau(\neg\mathrm{Con}_{\mathrm{PA}})) \vdash \neg\mathrm{Con}_{\mathrm{PA}}.$$

Hence by Cut we have

$$\mathrm{PA} \vdash \neg\mathrm{Con}_{\mathrm{PA}},$$

once again contradicting Gödel's second incompleteness theorem.

It follows by *reductio* that

$$\mathrm{HF},\ \tau(\mathrm{Con}_{\mathrm{PA}}) \nvdash \bot.$$

This completes the explanation of why $\tau(\mathrm{Con}_{\mathrm{PA}})$ can be neither proved nor refuted in first-order HF. (Interestingly, clauses (ii) and (iii) sufficed for the argument here.)

It now follows that the formula

$$x = x \wedge \tau(\mathrm{Con}_{\mathrm{PA}})$$

mentioned above could not count as *definit*, hence would be ineligible to serve as a separating formula. (Note also that first-order HF has non-standard models, just as first-order PA does. There is a non-standard model of HF that makes $\tau(\mathrm{Con}_{\mathrm{PA}})$ false.)

The limitative theorems mentioned above make all the more pressing the need for the constructivist to explicate *Definitheit* on behalf of Zermelo. Let us set aside the orderism issue, and rest content with working at first order. What are the right conditions for the constructivist to impose on the separating formulae that may be used in the Axiom Scheme of Separation?

It is worth asking whether there is a Brouwerian answer to this question, that can come to the rescue of the constructivist who wishes to espouse Choice but eschew Excluded Middle, in response to the Goodman–Myhill proof. A satisfactory explication of *Definitheit* will reveal that the separating formula at work in the Goodman–Myhill proof does *not* yield a legitimate instance of Separation. This is because in the actual 'separating formula' that Goodman and Myhill use:

$$x \in \{0\} \wedge \psi$$

it is the constituent ψ that features in the unwanted 'arbitrary instance' $\psi \vee \neg\psi$ of Excluded Middle that Goodman and Myhill visit upon the unsuspecting constructivist. Of course, $\psi \vee \neg\psi$ is perfectly acceptable to the constructivist provided that ψ is *effectively decidable*; otherwise, it is unacceptable.

That observation yields the following direct solution to the problem of how best to characterize *Definitheit*: *the separating formula should be effectively decidable, or already decided*. That completely draws the sting of the Goodman–Myhill result. For it then establishes Excluded Middle only for such formulae ψ as would yield instances $\psi \vee \neg\psi$ of Excluded Middle that are completely acceptable to the constructivist. Note that among the effectively decidable formulae are those that are actually *decided* — that is, *proved* or *refuted*. So Separation can be stated in two parts. One is the rule form discussed earlier:

Separation (first part)

$$\dfrac{\exists! t \qquad \begin{array}{c}\overline{a \in t}^{\,(i)} \\ \vdots \\ \varphi(a) \vee \neg\varphi(a)\end{array}}{\exists!\{x \in t|\varphi\}}{\scriptstyle (i)} \qquad a \text{ parametric.}$$

The second part of Separation would take the form

Separation (second part)

$$\dfrac{\exists! t}{\exists!\{x \in t|\varphi\}} \quad \text{provided } \Phi_t(\varphi),$$

where Φ_t is a decidable syntactic property of formulae φ that ensures their effective decidability on members of t in HF. An obvious starting candidate for Φ_t would be '... contains no unrestricted quantifiers'. There could be more inclusive properties Φ'_t of formulae that could in due course take over from such an initial choice for Φ_t; so the constructivist could treat Separation as an 'open-textured' part of their axiomatic basis. All that is important is that Φ'_t be a decidable syntactic property of formulae that recognizably guarantees the HF-decidability (on members of t) of any formula φ possessing Φ'_t.[32]

Because there is no way of knowing in advance what human ingenuity might serve up as instances of the subordinate (and constructive!) proof

$$\begin{array}{c} a \in t \\ \vdots \\ \varphi(a) \vee \neg\varphi(a) \end{array}$$

for potential application of (the first part of) Separation to obtain a subset of t using the separating formula $\varphi(x)$, one could well expect there to be such

[32] Thanks to an anonymous referee for pointing out that decidability of a separating formula *on members of t* is all that is needed. One can have t-relative decidability without having decidability on arbitrary members of the domain.

felicitously discovered formulae $\varphi(x)$ that do not, themselves, possess whatever decidable syntactic property Φ' is being applied in the second part of Separation just bruited. Despite this, the first part of Separation would be perfectly licit, courtesy of the *decision* rendered by the subordinate proof in hand. The constructivist would be meeting Zermelo's stronger requirement for *Definitheit*.

Let us now consider the question whether the second part of Separation stated above really does provide for greater 'power of proof' within (constructive) HF than would be enjoyed if one were to stick to the first part alone. The answer is negative. The only advantage of the second part is that it would allow one to reduce one's derivational burden with respect to the crucial subordinate proof. The second part of Separation would allow one simply to carry out an effective check on the Φ'-ness of φ, and proceed to apply Separation if indeed $\Phi'(\varphi)$ is found to be the case. But of course Φ' would have been adopted (for the second part of Separation) only after the theorist had determined that its application to formulae was indeed effectively decidable; and upon reaching the positive answer (regarding a formula φ) that $\Phi'(\varphi)$, one would then be in a position effectively to determine an appropriate subordinate (and constructive) proof of the kind required in order to be able to apply the *first part* of Separation. Such a proof, to be sure, might significantly increase the overall length of the proof that one would construct simply by helping oneself instead to the second part of Separation. But the point is, it would be discoverable in principle. One can therefore conclude that the second part of Separation does not properly extend the stock of theorems that the constructive set theorist would in principle be able to prove. It would only afford more efficient proofs for certain ones among them.

Having attained this clarity about a constructively acceptable form for Separation by reflecting on its role in HF, the constructivist can — and ought to — stick to this newly discovered form when venturing beyond HF and adopting the Axiom of Infinity in place of the Axiom of Finitude.

It would remain, then, to investigate exactly how much of (classical) ZFC the constructivist would be able to prove, by means of constructive logical reasoning, using the rules for set abstraction laid out above. One potentially fruitful avenue of inquiry would concern how, exactly, one should formulate the Power Set axiom

$$\forall x \exists w\, w = \{y | \forall z(z \in y \rightarrow z \in x)\}$$

or a rule of inference equivalent to it:

$$\frac{\exists! t}{\exists! \{y | \forall z(z \in y \rightarrow z \in t)\}}\ .$$

No longer would (or should) one be recognizing for membership in the power set of x all subsets of x *tout court*. Rather, one should recognize only those subsets of x that are *separable* from x (assuming one is given an oracle for membership in x). Thus the Axiom of Power Set would partake of the austerity imposed by the recommended reformulation of *Definitheit* on the part of separating formulae used in the Axiom of Separation.[33]

These investigations would require, however, another long paper, or even a monograph.

The author is offering here no theory of his own as to what pure sets, from the constructivist's point of view, should be postulated as existing. He is offering only a *logic* of sets, with no ontological commitments. That analytic basis can be shared by both classicists and constructivists. He is undecided as to whether the natural numbers should be assumed given (as *Urelemente*), or should be identified with, say, the finite von Neumann ordinals, should these be vouchsafed by future set-theoretic commitments. He is not (yet) proposing any definitional or constructive route to the real numbers by means of whatever sets he *might* well, in future, commit himself to. He is not certain whether the right road to the reals, for the constructivist, *is* by way of Bishop's classic book. And he is not certain that the right way for the constructivist to *found* mathematics is by copying what post-Cantorian classicists did, in trying to unify the various branches of mathematics in an extensional theory of *sets*. Perhaps there is some better and alternative way to provide a foundation for constructive mathematics, by treating the various kinds of numbers (natural, rational, real, complex, ...) as either logical objects, or objects *sui generis*, not to be identified with any sets.

Whatever alternative *set* theory might be proposed by a more enlightened constructivism, it will have to be proved (preferably, constructively!) that it does *not* admit a proof of Excluded Middle. Giving up one of the premises of the Goodman–Myhill proof is no guarantee, by itself, that there is no other deductive route to Excluded Middle from the remaining premises of their proof in combination with whatever other first principles are available in the new theory. The price of freedom (from Excluded Middle) is eternal vigilance.

REFERENCES

Aczel, Peter [1978]: 'The type theoretic interpretation of constructive set theory', in Angus Macintyre, Leszek Pacholski, and Jeff Paris, eds, *Logic Colloquium '77*, pp. 55–66. Amsterdam, New York, Oxford: North-Holland.

Aczel, Peter, and Michael Rathjen [2010]: *CST Book Draft*. Unpublished typescript, available for download at https://www1.maths.leeds.ac.uk/~rathjen/book.pdf. Accessed November 2019.

[33] There might be an interesting connection to be explored here with the contention of Eastaugh [2019] that the reversals of mathematical theorems in Reverse Mathematics 'track closure conditions' on the power set of the naturals. Since Reverse Mathematics has been developed, thus far, mainly for classical mathematical theorizing, it is difficult to estimate whether Eastaugh's contention would gain even more support from a more fully developed Constructive Reverse Mathematics. A place to start in such an investigation would be §5 of [Bridges and Palmgren, 2018]. One must note, however, the authors' 'stress that we restrict ourselves here to *informal* CRM, in which we take for granted the principles of function- and set-construction described in the first chapters of [Bishop, 1967] ...'.

Beeson, Michael J. [1979]: ‘Continuity in intuitionistic set theories’, in [Boffa *et al.*, 1979], pp. 1–52.

______ [1981]: ‘Formalizing constructive mathematics: Why and how?’, in [Richman, 1981], pp. 146–190.

______ [1982]: ‘Problematic principles in constructive mathematics’, in D. van Dalen, D. Lascar, and T. Smiley, eds, *Logic Colloquium ’80*, pp. 11–55. Amsterdam: North-Holland.

Bishop, Errett [1967]: *Foundations of Constructive Analysis*. New York: McGraw-Hill.

Boffa, Maurice, Dirk van Dalen, and Kenneth MacAloon, eds [1979]: *Logic Colloquium ’78*. Amsterdam: North-Holland.

Bridges, Douglas, and Erik Palmgren [2018]: ‘Constructive mathematics’, in Edward N. Zalta, ed., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University. Summer 2018 ed.

Brouwer, L.E.J. [1908]: ‘De onbetrouwbaarheid der logische principes’, *Tijdschrift voor wijsbegeerte* **2**, 152–158.

______ [1912]: *Intuitionisme en Formalisme*. Groningen: Noordhoff. English translation in *Bulletin of the American Mathematical Society* **20** (1913), 81–96.

______ [1918]: ‘Begründung der Mengenlehre unabhängig vom logischen Satz vom ausgeschlossenen Dritten. Erster Teil: Allgemeine Mengenlehre [Foundation of set theory independent of the logical law of the excluded middle. Part one: General set theory]’, *Verhandelingen der Koninklijke Akademie van Wetenschappen te Amsterdam, Eerste sectie.* **12**, No. 5, 1–43.

______ [1919]: ‘Intuitionistische Mengenlehre’, *Jahresbericht der Deutschen Mathematiker-Vereinigung* **28**, 203–208.

______ [1975]: ‘Intuitionism and formalism’, in A. Heyting, ed., *L.E.J. Brouwer: Collected Works 1: Philosophy and Foundations of Mathematics*, pp. 123–149. Amsterdam, New York, Oxford: North-Holland.

Dedekind, Richard [1888]: *Was sind und was sollen die Zahlen?* Braunschweig: Vieweg.

______ [1901]: *Essays on the Theory of Numbers*. Wooster Woodruff Beman, trans. Chicago: Open Court.

Eastaugh, Benedict [2019]: ‘Set existence principles and closure conditions: Unravelling the standard view of reverse mathematics’, *Philosophia Mathematica* (3) **27**, 153–176.

Ebbinghaus, Heinz-Dieter, Craig G. Fraser, and Akihiro Kanamori, eds [2010]: *Ernst Zermelo - Collected Works/Gesammelte Werke - Volume I/Band I - Set Theory, Miscellanea/Mengenlehre, Varia*. SpringerLink.
https://www.springer.com/gp/book/9783540793830.

Feferman, S. [1979]: ‘Constructive theories of functions and classes’, in [Boffa *et al.*, 1979], pp. 159–224.

Fraenkel, Adolf [1922]: ‘Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre’, *Mathematische Annalen* **86**, 230–237.

Friedman, Harvey M. [1973]: ‘The consistency of classical set theory relative to a set theory with intuitionistic logic’, *Journal of Symbolic Logic* **38**, 315–319.

______ [1977]: ‘Set theoretic foundations for constructive analysis’, *Annals of Mathematics* **105**, 1–28.

______ [2003]: *Vigre Lectures. Adventures in the Foundations of Mathematics.* ‘Lecture 2, The Axioms for Mathematics’. Unpublished typescript, pp. 18–32.
https://cpb-us-w2.wpmucdn.com/u.osu.edu/dist/1/1952/files/2014/01/
VIGRE050903-25g4l7w.pdf. Accessed February 2020.

Goodman, Nicholas D., and J. Myhill [1978]: 'Choice implies excluded middle', *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* **24**, 461.

Greenleaf, Newcomb [1981]: 'Liberal constructive set theory', in [Richman, 1981], pp. 213–240.

Henkin, Leon [1950]: 'Completeness in the theory of types', *The Journal of Symbolic Logic* **15**, 81–91.

Kaye, Richard, and Tim Lok Wong [2007]: 'On interpretations of arithmetic and set theory', *Notre Dame Journal of Formal Logic* **48**, 497–510.

Kronecker, L. [1886]: 'Ueber einige Anwendungen der Modulsysteme auf elementare algebraische Fragen', *Journal für die reine und angewandte Mathematik* **99**, 329–371.

McCarty, Charles [1986]: 'Realizability and recursive set theory', *Annals of Pure and Applied Logic* **32**, 153–183.

Myhill, J. [1975]: 'Constructive set theory', *The Journal of Symbolic Logic* **40**, 347–383.

Richman, Fred, ed. [1981]: *Constructive Mathematics*. Lecture Notes in Mathematics; 873. Springer-Verlag.

Shapiro, Stewart [1991]: *Foundations without Foundationalism: A Case for Second-Order Logic*. Oxford: Clarendon Press.

Tait, William [1994]: 'The law of excluded middle and the axiom of choice', in Alexander George, ed., *Mathematics and Mind*, pp. 45–70. Oxford University Press.

Tennant, Neil [2000]: 'Deductive versus expressive power: A pre-Gödelian predicament', *Journal of Philosophy* **97**, 257–277.

Troelstra, A.S., and D. van Dalen, [1988]: *Constructivism in Mathematics: An Introduction*. Vol. II. Amsterdam: Elsevier.

van Dalen, Dirk [2013]: *L.E.J. Brouwer – Topologist, Intuitionist, Philosopher*. Springer-Verlag.

van Heijenoort, Jean [1967]: *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*. Harvard University Press.

Zermelo, E. [1908]: 'Untersuchungen über die Grundlagen der Mengenlehre. I', *Mathematische Annalen* **65**, 261–281.

______ [1929]: 'Über den Begriff der Definitheit in der Axiomatik', *Fundamenta Mathematica* **14**, 339–344.