PREDICATIVITY, CIRCULARITY, AND ANTI-FOUNDATION

M. RATHJEN

Abstract. The anti-foundation axiom, **AFA**, has turned out to be a versatile principle in set theory for modelling a plethora of circular and self-referential phenomena. This paper explores whether **AFA** and the most important tools emanating from it, such as the solution lemma and the co-recursion principle, can be developed on predicative grounds, that is to say, within a predicative theory of sets.

If one could show that most of the circular phenomena that have arisen in computer science do not require impredicative set existence axioms for their modelling, this would demonstrate that their circularity is of a different kind than the one which underlies impredicative definitions.

1. Introduction

Russell discovered his paradox in May of 1901 while working on his *Principles of Mathematics* [28]. In response to the paradox he developed his distinction of logical types. Although first introduced in [28], type theory found its mature expression five years later in his 1908 article *Mathematical Logic as Based on the Theory of Types* [29]. In [29] Russell commences with a list of contradictions to be solved. These include Epimenides' liar paradox, his own paradox, and Burali-Forti's antinomy of the greatest ordinal. In a first analysis he remarks: *In all the above contradictions (which are merely selections from an indefinite number) there is a common characteristic, which we may describe as self-reference or reflexiveness.* ([29], p. 224). On closer scrutiny he discerns the form of reflexiveness that is the common underlying root for the trouble as follows:

> *Whatever we suppose to be the totality of propositions, statements about this totality generate new propositions which, on pain of contradiction, must lie outside the totality. It is useless to enlarge the totality, for that equally enlarges the scope of statements about the totality.* ([29], p. 224)

Here Russell declares very lucidly a ban on so-called *impredicative definitions*, first enunciated by Poincaré. An impredicative definition of an object refers to a presumed totality of which the object being defined is itself to be a member. For example, to define a set of natural numbers X as $X = \{n\in\mathbb{N} : \forall Y \subseteq \mathbb{N}\ F(n, Y)\}$ is impredicative since it involves the quantified variable 'Y' ranging over arbitrary

I am grateful to the Mittag-Leffler Institute (Djursholm, Sweden) for giving me the opportunity in 2001 to be working on the talk on which this paper is based. The research reported in this paper was also partly supported by United Kingdom Engineering and Physical Sciences Research Council Grant GR/R 15856/01.

subsets of the natural numbers $\mathbb{N}$, of which the set X being defined is one member. Determining whether $\forall Y \subseteq \mathbb{N}\, F(n, Y)\}$ holds involves an apparent circle since we shall have to know in particular whether $F(n, X)$ holds - but that cannot be settled until X itself is determined. Impredicative set definitions permeate the fabric of Zermelo-Fraenkel set theory in the guise of the separation and replacement axioms as well as the powerset axiom.

The avoidance of impredicative definitions has also been called the *Vicious Circle Principle*. This principle was taken very seriously by Hermann Weyl:

> *The deepest root of the trouble lies elsewhere: a field of possibilities open into infinity has been mistaken for a closed realm of things existing in themselves. As Brouwer pointed out, this is a fallacy, the Fall and Original Sin of set theory, even if no paradoxes result from it* ([34], p. 243).

When it turned out that the ramified theory of types rendered much of elementary mathematics unworkable, Russell, oblivious of his original insight, introduced, as an ad hoc device and for entirely pragmatic reasons, his notorious *axiom of reducibility*. This 'axiom' says that every set of higher level is coextensive with one of lowest level. Thereby he reconstituted the abolished impredicative definitions through the back door, as it were, undermining the the whole rationale behind a ramified hierarchy of types and levels to such an extent that it might just as well have been jettisoned altogether. Weyl derisively proclaimed:

> *Russell in order to extricate himself from the affair, causes reason to commit hara-kiri, by postulating the above assertion in spite of its lack of support by any evidence ('axiom of reducibility'). In a little book* Das Kontinuum, *I have tried to draw the honest consequence and constructed a field of real numbers of the first level, within which the most important operations of analysis can be carried out* ([34], p. 50).

In his analyzes of the contradictions and antinomies, Russell frequently equates the common culprit with self-reference (cf. above) and sees the problems as arising from *reflexive fallacies* (cf. [29], p. 230). He bans propositions of the form *"x is among x's"* (cf. [28], p. 105) and in his own so-called "no class" theory of classes he explains that propositional functions, such as the function *"x is a set"*, may not be applied to themselves as self-application would give rise to a vicious circle. Unlike his cogent analysis of the problem of impredicativity, Russell's charges against self-referential notions are not that well explained. It might be that the scope of his criticism was just intended to be confined to questions of concern to his own type theories, though it is more likely that he regarded self-referential notions as being incoherent or rather fallacious conceptions that should be banned all the same. In the history of philosophy, the charge of circularity is old and regarded as tantamount to delivering a refutation. In his *Analytica Posteriora*, Aristotle outlaws circular lines of arguments. Similarly, Aquinas calls an infinite series of reasons each of which is in some sense dependent on a prior a *vicious regress*. On the other hand, in hermeneutics it has been a tenet that comprehension can only come about through a tacit foreknowledge, thereby emphasizing the inherent circularity of all understanding (the *hermeneutic circle*) . In the same vein, circularity has been found in much intentional activity related to self-consciousness, communication, common knowledge, and conventions. For instance, on D. Lewis's account in his book *Convention* [20], for something to be a convention among a group of people, common knowledge about certain facts must obtain in this group. If C and E are modal operators such that $C\varphi$ and $E\varphi$ stand for *'φ is common knowledge'* and *'φ is known to everybody'*, respectively, then the central axiom which implicitly defines C takes the self-referential form $C\varphi \leftrightarrow E(\varphi \wedge C\varphi)$.

In logic, circular concepts involving self-reference or self-application have proven to be very important, as witnessed by Gödel's incompleteness theorems, the recursion theorem in recursion theory, self-applicative programs, and such so-called applicative theories as Feferman's *Explicit Mathematics* that have built-in gadgets allowing for self-application.

The general blame that Russell laid on circularity was more influential then the more specific ban placed on impredicative definitions. In the wake of Russell's anathematizing of circularity, Tarski's hierarchical approach via meta-languages became the accepted wisdom on the semantical paradoxes. On the other hand, the fruitfulness of circular notions in many areas of scientific discourse demonstrated that these notions are scientifically important. Thus one is naturally led to search for criteria that would enable one to tell the benign (and fruitful) circularities from the paradoxical ones. Kripke, though, in his *Outline of a Theory of Truth* [19], demonstrated rather convincingly that there are no such general 'syntactic' criteria for making this distinction, in that whether or not something is paradoxical may well depend on non-linguistic facts.

Notwithstanding the lack of simple syntactical criteria for detecting paradoxical circularities, it is of great interest to develop frameworks in which many non-paradoxical circular phenomena can be modelled. One such framework is Feferman's theory of explicit mathematics (cf. [13]). It is suitable for representing Bishop-style constructive mathematics as well as generalized recursion, including direct expression of structural concepts which admit self-application. Another very systematic toolbox for building models of various circular phenomena is set theory with the *Anti-Foundation axiom.* Theories like **ZF** outlaw sets like $\Omega = \{\Omega\}$ and infinite chains of the form $\Omega_{i+1} \in \Omega_i$ for all $i \in \omega$ on account of the Foundation axiom, and sometimes one hears the mistaken opinion that the only coherent conception of sets precludes such sets. The fundamental distinction between well-founded and non-well-founded sets was formulated by Mirimanoff in 1917. The relative independence of the Foundation axiom from the other axioms of Zermelo-Fraenkel set theory was announced by Bernays in 1941 but did not appear until the 1950s. Versions of axioms asserting the existence of non-well-founded sets were proposed by Finsler (1926). The ideas of Bernays' independence proof were exploited by Rieger, Hájek, Boffa, and Felgner. After Finsler, Scott in 1960 appears to have been the first person to consider an anti-foundation axiom which encapsulates a strengthening of the axiom of extensionality. The anti-foundation axiom in its strongest version was first formulated by Forti and Honsell [16] in 1983. Though several logicians explored set theories whose universes contained non-wellfounded sets (or hypersets as they are called nowadays) the area was considered rather exotic until these theories were put to use in developing rigorous accounts of circular notions in computer science (cf. [3]). It turned out that the *Anti-Foundation Axiom,* **AFA**, gave rise to a rich universe of sets and provided an elegant tool for modelling all sorts of circular phenomena. The application areas range from modal logic, knowledge representation and theoretical economics to the semantics of natural language and programming languages. The subject of hypersets and their applications is thoroughly developed in the books [3] by P. Aczel and [5] by J. Barwise and L. Moss.

While reading [3] and [5], the question that arose in my mind was that of whether or not the material could be developed on the basis of a constructive universe of hypersets rather than a classical and impredicative one. This paper explores whether **AFA** and the most important tools emanating from it, such as the solution lemma and the co-recursion principle, can be developed on predicative grounds, that is to say, within a predicative theory of sets. The upshot will be that most of the circular phenomena that have arisen in computer science don't require impredicative set existence axioms for their modelling, thereby showing that their circularity is clearly of a different kind than the one which underlies impredicative definitions.

2. The anti-foundation axiom

Definition 2.1. A *graph* will consist of a set of *nodes* and a set of *edges*, each edge being an ordered pair $\langle x, y\rangle$ of nodes. If $\langle x, y\rangle$ is an edge then we will write $x \to y$ and say that y is a *child* of x.

A *path* is a finite or infinite sequence $x_0 \to x_1 \to x_2 \to \ldots$ of nodes $x_0, x_1, x_2, \ldots$ linked by edges $\langle x_0, x_1\rangle, \langle x_1, x_2\rangle, \ldots$.

A *pointed graph* is a graph together with a distinguished node x_0 called its *point.* A pointed graph is *accessible* if for every node x there is a path $x_0 \to x_1 \to x_2 \to \ldots \to x$ from the point x_0 to x.

A *decoration* of a graph is an assignment d of a set to each node of the graph in such a way that the elements of the set assigned to a node are the sets assigned to the children of that node, i.e.

$$d(a) = \{d(x) : \ a \to x\}.$$

A *picture* of a set is an accessible pointed graph (apg for short) which has a decoration in which the set is assigned to the point.

Definition 2.2. The *Anti-Foundation Axiom*, **AFA**, is the statement that every graph has a unique decoration.

Note that **AFA** has the consequence that every apg is a picture of a unique set.

AFA is in effect the conjunction of two statements:

- $\mathbf{AFA}_1$: *Every graph has at least one decoration.*
- $\mathbf{AFA}_2$: *Every graph has a most one decoration.*

$\mathbf{AFA}_1$ is an existence statement whereas $\mathbf{AFA}_2$ is a strengthening of the Extensionality axiom of set theory. For example, taking the graph $\mathbb{G}_0$ to consist of a single node x_0 and one edge $x_0 \to x_0$, $\mathbf{AFA}_1$ ensures that this graph has a decoration $d_0(x) = \{d_0(y) : x \to y\} = \{d_0(x)\}$, giving rise to a set b such that $b = \{b\}$. However, if there is another set c satisfying $c = \{c\}$, the Extensionality axiom does not force b to be equal to c, while $\mathbf{AFA}_2$ yields $b = c$. Thus, by **AFA** there is exactly one set Ω such that $\Omega = \{\Omega\}$.

Another example which demonstrates the extensionalizing effect of $\mathbf{AFA}_2$ is provided by the graph $\mathbb{G}_\infty$ which consists of the infinitely many nodes x_i and the edges $x_i \to x_{i+1}$ for each $i \in \omega$. According to $\mathbf{AFA}_1$, $\mathbb{G}_\infty$ has a decoration. As $d_\infty(x_i) = \Omega$ defines such a decoration, $\mathbf{AFA}_2$ entails that this is the only one, whereby the different graphs $\mathbb{G}_0$ and $\mathbb{G}_\infty$ give rise to the same non-well-founded set.

The most important applications of **AFA** arise in connection with solving systems of equations of sets. In a nutshell, this is demonstrated by the following example. Let p and q be arbitrary fixed sets. Suppose we need sets x, y, z such that

$$\begin{aligned} x &= \{x, y\} \\ y &= \{p, q, y, z\} \\ z &= \{p, x, y\}. \end{aligned} \tag{1}$$

Here p and q are best viewed as atoms while x, y, z are the indeterminates of the system. **AFA** ensures that the system (1) has a unique solution. There is a powerful technique that can be used to show that systems of equations of a certain type have always unique solutions. In the terminology of [5] this is called the *solution lemma*. We shall prove it in the sections on applications of **AFA**.

3. **AFA** in constructive set theory

In this section I will present some results about the proof-theoretic strength of systems of constructive set theory with **AFA** instead of $\in$-Induction.

3.1. **Constructive set theory.** Constructive set theory grew out of Myhill's endeavours (cf. [23]) to discover a simple formalism that relates to Bishop's constructive mathematics as **ZFC** relates to classical Cantorian mathematics. Later on Aczel modified Myhill's set theory to a system which he called Constructive Zermelo-Fraenkel set theory, **CZF**, and corroborated its constructiveness by interpreting it in Martin-Löf type theory (**MLTT**) (cf. [1]). The interpretation was in many ways canonical and can be seen as providing **CZF** with a standard model in type theory.

Let $\mathbf{CZF}^-$ be **CZF** without $\in$-induction and let **CZFA** be $\mathbf{CZF}^-$ plus **AFA**. I. Lindström (cf. [21]) showed that **CZFA** can be interpreted in **MLTT** as well. Among other sources, the work of [21] will be utilized in calibrating the exact strength of various extensions of **CZFA**, in particular ones with inaccessible set axioms. The upshot is that **AFA** does not yield any extra proof-theoretic strength on the basis of constructive set theory and is indeed much weaker in proof strength than $\in$-Induction. This contrasts with Kripke-Platek set theory, **KP**. The theory **KPA**, which adopts **AFA** in place of the Foundation Axiom scheme, is proof-theoretically considerably stronger than **KP** as was shown in [26]. On the other hand, while being weaker in proof-theoretic strength, **CZFA** seems to be "mathematically" stronger than **KPA** in that most applications that **AFA** has found can be easily formalized in **CZFA** whereas there are serious difficulties with doing this in **KPA**. For instance, the proof from **AFA** that the collection of streams over a given set A exists and forms a set, seems to require the exponentiation axiom, a tool which is clearly not available in **KPA**.

3.2. **The theory CZFA.** The language of **CZF** is the first order language of Zermelo-Fraenkel set theory, LST, with the non logical primitive symbol $\in$. We assume that LST has also a constant, ω, for the set of the natural numbers.

Definition 3.1 (Axioms of **CZF**)**. CZF** is based on intuitionistic predicate logic with equality. The set theoretic axioms of **CZF** are the following:

(1) **Extensionality** $\forall a\,\forall b\,(\forall y\,(y \in a \leftrightarrow y \in b) \rightarrow a = b)$.
(2) **Pair** $\forall a\,\forall b\,\exists x\,\forall y\,(y \in x \leftrightarrow y = a \vee y = b)$.
(3) **Union** $\forall a\,\exists x\,\forall y\,(y \in x \leftrightarrow \exists z{\in}a\, y \in z)$.
(4) Δ_0 **- Separation scheme** $\forall a\,\exists x\,\forall y\,(y \in x \leftrightarrow y \in a \wedge \varphi(y))$,
for every *bounded* formula $\varphi(y)$, where a formula $\varphi(x)$ is bounded, or Δ_0, if all the quantifiers occurring in it are bounded, i.e. of the form $\forall x{\in}b$ or $\exists x{\in}b$.
(5) **Subset Collection scheme**

$$\begin{aligned}&\forall a\,\forall b\,\exists c\,\forall u\,\big(\forall x{\in}a\,\exists y{\in}b\;\varphi(x,y,u) \rightarrow\\ &\exists d{\in}c\,(\forall x{\in}a\,\exists y{\in}d\;\varphi(x,y,u) \wedge \forall y{\in}d\,\exists x{\in}a\;\varphi(x,y,u))\big)\end{aligned}$$

for every formula $\varphi(x,y,u)$.
(6) **Strong Collection scheme**

$$\begin{aligned}&\forall a\,\big(\forall x{\in}a\,\exists y\;\varphi(x,y) \rightarrow\\ &\exists b\,(\forall x{\in}a\,\exists y{\in}b\;\varphi(x,y) \wedge \forall y{\in}b\,\exists x{\in}a\;\varphi(x,y))\big)\end{aligned}$$

for every formula $\varphi(x,y)$.
(7) **Infinity**

$$\begin{array}{ll}(\omega 1) & 0{\in}\omega \wedge \forall y\,(y{\in}\omega \rightarrow y+1{\in}\omega)\\ (\omega 2) & \forall x\,(0{\in}x \wedge \forall y\,(y{\in}x \rightarrow y+1{\in}x) \rightarrow \omega{\subseteq}x),\end{array}$$

where $y+1$ is $y \cup \{y\}$, and 0 is the empty set, defined in the obvious way.
(8) $\in$ **- Induction scheme**

$$(IND_{\in}) \qquad \forall a\,(\forall x{\in}a\;\varphi(x) \rightarrow \varphi(a)) \rightarrow \forall a\;\varphi(a),$$

for every formula $\varphi(a)$.

Definition 3.2. Let $\mathbf{CZF}^-$ be the system **CZF** without the $\in$ - Induction scheme.

Remark 3.3. $\mathbf{CZF}^-$ is strong enough to show the existence of any primitive recursive function on ω and therefore Heyting arithmetic can be interpreted in $\mathbf{CZF}^-$ in the obvious way. By way of example, let's verify this for addition: As a consequence of Subset Collection one obtains that for arbitrary sets a, b, the class of all functions from a to b, ${}^ab$, is a set. Using Strong Collection, $\{{}^n\omega : n{\in}\omega\}$ is a set, and thus $\mathrm{Fin} := \bigcup_{n\in\omega} {}^{n}\omega$ is a set, too. Employing the axiom ($\omega2$) one shows that

$$\forall\langle n,m\rangle \in \omega\times\omega\, \exists! f \in \mathrm{Fin}\, \theta(n,m,f),$$

where $\theta(n,m,f)$ stands for the formula

$$\mathbf{dom}(f) = m+1 \ \wedge\ f(0) = n \ \wedge\ (\forall i{\in}m) f(i+1) = f(i)+1.$$

Using Strong Collection, there exists a set A such that

$$\forall\langle n,m\rangle \in \omega\times\omega\, \exists f \in A\, \theta(n,m,f) \ \wedge\ \forall f{\in}A\, \exists\langle n,m\rangle \in \omega\times\omega\, \theta(n,m,f).$$

Now define $h : \omega\times\omega \to \omega$ by letting $h(n,m) = k$ if and only if

$$\exists f \in A \left[\theta(n,m,f) \ \wedge\ f(m) = k\right].$$

It is easy to show that h satisfies the recursion equations

$$h(n,0) = n \quad\wedge\quad h(n,m+1) = h(n,m)+1.$$

Definition 3.4. Unfortunately, $\mathbf{CZF}^-$ has certain defects from a mathematical point of view in that this theory appears to be too limited for proving proving the existence of the transitive closure of an arbitrary set. To remedy this we shall consider an axiom, TRANS, which ensures that every set is contained in a transitive set:

$$\text{TRANS} \qquad\qquad \forall x\, \exists y \left[x \subseteq y \ \wedge\ (\forall u{\in}y)\,(\forall v{\in}u)\, v{\in}y\right].$$

Let **CZFA** be the theory $\mathbf{CZF}^- + \text{TRANS} + \mathbf{AFA}$.

Lemma 3.5. *Let* $\mathbf{TC}(x)$ *stand for the smallest transitive set that contains all elements of* x*.* $\mathbf{CZF}^- + \text{TRANS}$ *proves the existence of* $\mathbf{TC}(x)$ *for any set* x*.*

Proof: We shall use a consequence of Subset Collection called *Exponentiation* which asserts that for arbitrary sets a,b, the class of all functions from a to b, ${}^{a}b$, is a set.

Let x be an arbitrary set. By TRANS there exists a transitive set A such that $x \subseteq A$. For $n{\in}\omega$ let

$$\begin{aligned} B_n &= \{f \in {}^{n+1}A :\ f(0) \in x \ \wedge\ (\forall i{\in}n)\, f(i+1) \in f(i)\}, \\ \mathbf{TC}_n(x) &= \bigcup\{\mathbf{ran}(f) :\ f \in B_n\}, \end{aligned}$$

where $\mathbf{ran}(f)$ denotes the range of a function f. B_n is a set owing to Exponentiation and Δ_0 Separation. $\mathbf{TC}_n(x)$ is a set by Union. Furthermore, $C = \bigcup_{n\in\omega} \mathbf{TC}_n(x)$ is a set by Strong Collection and Union. Then $x = \mathbf{TC}_0(x) \subseteq C$. Let y be a transitive set such that $x \subseteq y$. By induction on n one easily verifies that $\mathbf{TC}_n(x) \subseteq y$, and hence $C \subseteq y$. Moreover, C is transitive. Thus C is the smallest transitive set which contains all elements of x. □

Definition 3.6. A mathematically very useful axiom to have in set theory is the *Dependent Choices Axiom*, **DC**, i.e., for all formulae ψ, whenever

$$(\forall x{\in}a)\,(\exists y{\in}a)\,\psi(x,y)$$

and $b_0 \in a$, then there exists a function $f : \omega \to a$ such that $f(0) = b_0$ and

$$(\forall n \in \omega)\, \psi(f(n), f(n+1)).$$

For a function f let $\mathbf{dom}(f)$ denote the domain of f. Even more useful in constructive set theory is the *Relativized Dependent Choices Axiom*, **RDC**.[1] It asserts that for arbitrary formulae ϕ and ψ, whenever

$$\forall x\big[\phi(x) \;\to\; \exists y\big(\phi(y) \;\wedge\; \psi(x,y)\big)\big]$$

and $\phi(b_0)$, then there exists a function f with $\mathbf{dom}(f) = \omega$ such that $f(0) = b_0$ and

$$(\forall n \in \omega)\big[\phi(f(n)) \;\wedge\; \psi(f(n), f(n+1))\big].$$

A restricted form of **RDC** where ϕ and ψ are required to be Δ_0 formulas will be called Δ_0-**RDC**.

The defect of $\mathbf{CZF}^-$ concerning the lack of enough transitive sets can also be remedied by adding Δ_0-**RDC** to $\mathbf{CZF}^-$. It is perhaps worth noting that Δ_0-**RDC** implies **DC** on the basis of $\mathbf{CZF}^-$.

Lemma 3.7. $\mathbf{CZF}^- + \Delta_0$-$\mathbf{RDC} \vdash \mathbf{DC}$.

Proof: See [27], Lemma 3.4. □

The existence of the transitive closure of any set can also be obtained by slightly strengthening induction on ω to

$$\Sigma_1\text{-}\mathbf{IND}_\omega \quad \phi(0) \;\wedge\; (\forall n \in \omega)(\phi(n) \to \phi(n+1)) \;\to\; (\forall n \in \omega)\phi(n)$$

for all Σ_1 formulae ϕ. It is worth noting that Σ_1-$\mathbf{IND}_\omega$ actually implies

$$\Sigma\text{-}\mathbf{IND}_\omega \quad \theta(0) \;\wedge\; (\forall n \in \omega)(\theta(n) \to \theta(n+1)) \;\to\; (\forall n \in \omega)\theta(n)$$

for all Σ formulae θ, where the Σ formulae are the smallest collection of formulae comprising the Δ_0 formulae which is closed under $\wedge, \vee$, bounded quantification, and (unbounded) existential quantification. This is due to the fact that every Σ formula is equivalent to a Σ_1 formula provably in $\mathbf{CZF}^-$. The latter principle is sometimes called the Σ *Reflection Principle* and can be proved as in Kripke-Platek set theory (one easily verifies that the proof of [4], I.4.3 also works in $\mathbf{CZF}^-$). Σ-$\mathbf{IND}_\omega$ enables one to introduce functions by Σ recursion on ω (cf. [4], I.6) as well as the transitive closure of an arbitrary set (on the basis of $\mathbf{CZF}^-$). It is worth noting that Σ-$\mathbf{IND}_\omega$ is actually a consequence of Δ_0-**RDC**.

Lemma 3.8. $\mathbf{CZF}^- + \Delta_0$-$\mathbf{RDC} \vdash \Sigma$-$\mathbf{IND}_\omega$.

Proof: Suppose $\theta(0) \;\wedge\; (\forall n \in \omega)(\theta(n) \;\to\; \theta(n+1))$, where $\theta(n)$ is of the form $\exists x \phi(n,x)$ with ϕ Δ_0. We wish to prove $(\forall n \in \omega)\theta(n)$.

If z is an ordered pair $\langle x, y\rangle$ let $1^{st}(z)$ denote x and $2^{nd}(z)$ denote y. Since $\theta(0)$ there exists a set x_0 such that $\phi(0, x_0)$. Put $a_0 = \langle 0, x_0\rangle$.

[1] In Aczel [2], **RDC** is called the dependent choices axiom and **DC** is dubbed the axiom of limited dependent choices. We deviate from the notation in [2] as it deviates from the usage in classical set theory texts.

From $(\forall n\in\omega)(\theta(n) \to \theta(n+1))$ we can conclude

$$(\forall n\in\omega)\forall y\,[\phi(n,y) \to \exists w\,\phi(n+1,w)\,]$$

and thus

$$\forall z\,\big[\,\psi(z) \to \exists v\,(\,\psi(v)\ \wedge\ \chi(z,v)\,)\big],$$

where $\psi(z)$ stands for z *is an ordered pair* $\wedge\ 1^{st}(z) \in \omega\ \wedge\ \phi(1^{st}(z), 2^{nd}(z))$ and $\chi(z,v)$ stands for $1^{st}(v) = 1^{st}(z) + 1$. Note that ψ and χ are Δ_0. We also have $\psi(a_0)$. Thus by Δ_0-**RDC** there exists a function $f : \omega \to V$ such that $f(0) = a_0$ and

$$(\forall n\in\omega)\,\big[\,\psi(f(n))\ \wedge\ \chi(f(n), f(n+1))\,\big].$$

From $\chi(f(n), f(n+1))$, using induction on ω, one easily deduces that $1^{st}(f(n)) = n$ for all $n\in\omega$. Hence from $(\forall n\in\omega)\,\psi(f(n))$ we get $(\forall n\in\omega)\,\exists x\,\phi(n,x)$ and so $(\forall n\in\omega)\,\theta(n)$. $\square$

We shall consider also the full scheme of induction on ω,

$$\mathbf{IND}_\omega \quad \psi(0)\ \wedge\ (\forall n\in\omega)(\psi(n) \to \psi(n+1))\ \to\ (\forall n\in\omega)\psi(n)$$

for all formulae ψ.

Lemma 3.9. $\mathbf{CZF}^- + \mathbf{RDC} \vdash \mathbf{IND}_\omega$.

Proof: Suppose $\theta(0) \wedge (\forall n\in\omega)(\theta(n) \to \theta(n+1))$. We wish to prove $(\forall n\in\omega)\theta(n)$. Let $\phi(x)$ and $\psi(x,y)$ be the formulas $x\in\omega \wedge \theta(x)$ and $y = x+1$, respectively. Then $\forall x\,[\phi(x) \to \exists y\,(\phi(y) \wedge \psi(x,y))]$ and $\phi(0)$. Hence, by **RDC**, there exists a function f with domain ω such that $f(0) = 0$ and $\forall n\in\omega\,[\phi(f(n))\ \wedge\ \psi(f(n), f(n+1))]$. Let $a = \{n\in\omega :\ f(n) = n\}$. Using the induction principle (ω2) one easily verifies $\omega \subseteq a$, and hence $f(n) = n$ for all $n\in\omega$. Hence, $\phi(n)$ for all $n\in\omega$, and thus $(\forall n\in\omega)\theta(n)$. $\square$

4. Predicativism

Weyl rejected the platonist philosophy of mathematics as manifested in impredicative existence principles of Zermelo-Fraenkel set theory. In his book *Das Kontinuum*, he initiated a predicative approach to the the real numbers and gave a viable account of a substantial chunk of analysis. What are the ideas and principles upon which his "predicative view" is supposed to be based? A central tenet is that there is a fundamental difference between our understanding of the concept of natural numbers and our understanding of the set concept. Like the French predicativists, Weyl accepts the completed infinite system of natural numbers as a point of departure. He also accepts classical logic but just works with sets that are of level one in Russell's ramified hierarchy, in other words only with the principle of arithmetical definitions.

Logicians such as Wang, Lorenzen, Schütte, and Feferman then proposed a foundation of mathematics using layered formalisms based on the idea of predicativity which ventured into higher levels of the ramified hierarchy. The idea of an autonomous progression of theories $RA_0, RA_1, \ldots, RA_\alpha, \ldots$ was first presented in Kreisel [18] and than taken up by Feferman and Schütte to determine the limits of predicativity. The notion of autonomy therein is based on introspection and should perhaps be viewed as a 'boot-strap' condition. One takes the structure of natural numbers as one's point of departure and then explores through a process of active reflection what is implicit in accepting this structure, thereby developing a growing body of ever higher layers of the ramified hierarchy. Feferman and Schütte (cf. [30, 31, 11, 12]) showed that the ordinal Γ_0 is the first ordinal whose well-foundedness cannot be proved in autonomous progressions of theories. It was also argued by Feferman that the whole sequence of autonomous progressions of theories is coextensive with predicativity, and on these grounds Γ_0 is often referred to as the proper limit of all predicatively provable ordinals. In this paper I shall only employ the "lower bound" part of this analysis, i.e., that every ordinal less than Γ_0 is a predicatively provable ordinal. In consequence, every theory with proof-theoretic ordinal less than Γ_0 has a predicative consistency proof and is moreover conservative over a theory RA_α for arithmetical statements for some $\alpha < \Gamma_0$. As a shorthand for the above I shall say that a theory is *predicatively justifiable.* The remainder of this section lists results showing that **CZFA** and its variants are indeed predicatively justifiable.

As a scale for measuring the proof-theoretic strength of theories one uses traditionally certain subsystems of second order arithmetic (see [14, 33]). Relevant to the present context are systems based on the Σ^1_1 axiom of choice and the Σ^1_1 axiom of dependent choices. The theory $\mathbf{\Sigma^1_1}$-**AC** is a subsystem of second order arithmetic with the Σ^1_1 axiom of choice and induction over the natural numbers for all formulas while $\mathbf{\Sigma^1_1}$-$\mathbf{DC_0}$ is a subsystem of second order arithmetic with the Σ^1_1 axiom of dependent choices and induction over the natural numbers restricted to formulas without second order quantifiers (for precise definitions see [14, 33]). The proof theoretic ordinal of $\mathbf{\Sigma^1_1}$-**AC** is $\varphi\varepsilon_0 0$ while $\mathbf{\Sigma^1_1}$-$\mathbf{DC_0}$ has the smaller proof-theoretic ordinal $\varphi\omega 0$ as was shown by Cantini [7]. Here φ denotes the Veblen function (see [32]).

Theorem 4.1. (i) *The theories* $\mathbf{CZF}^-+\Sigma_1$-$\mathbf{IND}_\omega$, $\mathbf{CZFA}+\Sigma_1\mathbf{IND}_\omega+\Delta_0$-**RDC**, **CZFA** $+ \Sigma_1$-$\mathbf{IND}_\omega$ + **DC**, *and* $\mathbf{\Sigma^1_1}$-$\mathbf{DC_0}$ *are proof-theoretically equivalent. Their proof-theoretic ordinal is* $\varphi\omega 0$.

(ii) *The theories* $\mathbf{CZF}^- + \mathbf{IND}_\omega$, **CZFA** + $\mathbf{IND}_\omega$ + **RDC**, $\widehat{\mathbf{ID}}_\mathbf{1}$, *and* $\mathbf{\Sigma^1_1}$-**AC** *are proof-theoretically equivalent. Their proof-theoretic ordinal is* $\varphi\varepsilon_0 0$.

(iii) **CZFA** *has at least proof-theoretic strength of Peano arithmetic and so its proof-theoretic ordinal is at least* ε_0. *An upper bound for the proof-theoretic ordinal of* **CZFA** *is* $\varphi 20$. *In consequence,* **CZFA** *is proof-theoretically weaker than* **CZFA** + Δ_0-**RDC**.

iv *All the foregoing theories are predicatively justifiable.*

Proof: (ii) follows from [27], Theorem 3.15.

As to (i) it is important to notice that the scheme dubbed Δ_0-**RDC** in [27] is not the same as Δ_0-**RDC** in the present paper. In [27], Δ_0-**RDC** asserts for Δ_0 formulas ϕ and ψ that whenever $(\forall x{\in}a)\big[\phi(x) \rightarrow (\exists y{\in}a)\big(\phi(y) \wedge \psi(x,y)\big)\big]$ and $b_0{\in}a \wedge \phi(b_0)$, then there exists a function $f : \omega \rightarrow a$ such that $f(0) = b_0$ and $(\forall n{\in}\omega)\big[\phi(f(n)) \wedge \psi(f(n), f(n+1))\big]$. The latter principle is weaker than our Δ_0-**RDC** as all quantifiers have to be restricted to a given set a. However, the realizability interpretation of constructive set theory in $\mathbf{PA}^r_\Omega + \Sigma^\Omega$-**IND** employed in the proof of [27], Theorem 3.15 (i) also validates the stronger Δ_0-**RDC** of the present paper (the system $\mathbf{PA}^r_\Omega$ stems from [17]).

Theorem 3.15 (i) of [27] and Lemma 3.8 also imply that $\mathbf{CZF}^- + \Delta_0$-**RDC** is not weaker than $\mathbf{CZF}^- + \Sigma_1$-$\mathbf{IND}_\omega$. Thus proof-theoretic equivalence of all systems in (i) ensues.

(iii) is a consequence of remark 3.3. At present the exact proof-theoretic strength of **CZFA** is not known, however, it can be shown that the proof-theoretic ordinal of **CZFA** is not bigger than $\varphi 20$. The latter bound can be obtained by inspecting the interpretation of **CZFA** in $\mathbf{PA}^r_\Omega + \Sigma^\Omega$-**IND** employed in the proof of [27], Theorem 3.15. A careful inspection reveals that a subtheory T of $\mathbf{PA}^r_\Omega + \Sigma^\Omega$-**IND** suffices. To be more precise, T can be taken to be the theory

$$\mathbf{PA}^r_\Omega + \forall\alpha\, \exists\lambda\, [\alpha < \lambda \ \wedge \ \lambda \text{ is a limit ordinal}].$$

Using cut elimination techniques and asymmetric interpretation, T can be partially interpreted in $\mathbf{RA}_{<\omega^2}$. The latter theory is known to have proof-theoretic ordinal $\varphi 20$.

(iv) The above ordinals are less than Γ_0. □

Remark 4.2. Constructive set theory with **AFA** has an interpretation in Martin-Löf type theory as has been shown by I. Lindström [21]. Martin-Löf type theory is considered to be the most acceptable foundational framework of ideas that makes precise the constructive approach to mathematics. The interpretation of **CZFA** in Martin-Löf type theory demonstrates that there is a constructive notion of set that lends constructive meaning to **AFA**. However, Martin-Löf type theory is not a predicative theory in the sense of Feferman and Schütte as it possesses a proof-theoretic ordinal bigger than Γ_0. The work in [27] shows that **CZFA** and its variants can also be reduced to theories which are predicative in the stricter sense of autonomous progressions.

5. On using the anti-foundation axiom

In this section I rummage through several applications of **AFA** made in [3] and [5]. In order to corroborate my claim that most applications of **AFA** require only constructive means, various sections of [3] and [5] are recast on the basis of the theory **CZFA** rather than **ZFA**.

5.1. **The Labelled Anti-Foundation Axiom.** In applications it is often useful to have a more general form of **AFA** at one's disposal.

Definition 5.1. A *labelled graph* is a graph together with a labelling function ℓ which assigns a set $\ell(a)$ of *labels* to each node a.

A *labelled decoration* of a labelled graph is a function d such that

$$d(a) \quad = \quad \{d(b) \, : \, a \to b\} \cup \ell(a).$$

An unlabelled graph $(G, \twoheadrightarrow)$ may be identified with the special labelled graph where the labelling function $\ell : G \to V$ always assigns the empty set, i.e. $\ell(x) = \emptyset$ for all $x \in G$.

Theorem 5.2. (**CZFA**) (Cf. [3], Theorem 1.9) *Each labelled graph has a unique labelled decoration.*

Proof: Let $\mathbb{G} = (G, \twoheadrightarrow, \ell)$ be a labelled graph. Let $\mathbb{G}' = (G', \to)$ be the graph having as nodes all the ordered pairs $\langle i, a\rangle$ such that either $i = 1$ and $a \in G$ or $i = 2$ and $a \in \mathbf{TC}(G)$ and having as edges:

- $\langle 1, a\rangle \to \langle 1, b\rangle$ whenever $a \twoheadrightarrow b$,
- $\langle 1, a\rangle \to \langle 2, b\rangle$ whenever $a \in G$ and $b \in \ell(a)$,
- $\langle 2, a\rangle \to \langle 2, b\rangle$ whenever $b \in a \in \mathbf{TC}(G)$.

By **AFA**, $\mathbb{G}'$ has a unique decoration π. So for each $a \in G$

$$\pi(\langle 1, a\rangle) \; = \; \{\pi(\langle 1, b\rangle) \, : \, a \twoheadrightarrow b\} \; \cup \; \{\pi(\langle 2, b\rangle) \, : \, b \in \ell(a)\}$$

and for each $a \in \mathbf{TC}(G)$,

$$\pi(\langle 2, a\rangle) \; = \; \{\pi(\langle 2, b\rangle) \, : \, b \in a\}.$$

Note that the set $\mathbf{TC}(G)$ is naturally equipped with a graph structure by letting its edges $x \multimap y$ be defined by $y \in x$. The unique decoration for $(\mathbf{TC}(G), \multimap)$ is obviously the identity function on $\mathbf{TC}(G)$. As $x \mapsto \pi(\langle 2, x\rangle)$ is also a decoration of $(\mathbf{TC}(G), \multimap)$ we can conclude that $\pi(\langle 2, x\rangle) = x$ holds for all $x \in \mathbf{TC}(G)$. Hence if we let $\tau(a) = \pi(\langle 1, a\rangle)$ for $a \in G$ then, for $a \in G$,

$$\tau(a) \; = \; \{\tau(b) \, : \, a \twoheadrightarrow b\} \; \cup \; \ell(a),$$

so that τ is a labelled decoration of the labelled graph $\mathbb{G}$.

For the uniqueness of τ suppose that τ' is a labelled decoration of $\mathbb{G}$. Then π' is a decoration of the graph $\mathbb{G}'$, where

$$\begin{aligned} \pi'(\langle 1, a\rangle) \quad &= \quad \tau'(a) \text{ for } a \in G, \\ \pi'(\langle 2, a\rangle) \quad &= \quad a \text{ for } a \in \mathbf{TC}(G). \end{aligned}$$

It follows from **AFA** that $\pi' = \pi$ so that for $a \in G$

$$\tau'(a) = \pi'(\langle 1, a\rangle) = \pi(\langle 1, a\rangle) = \tau(a),$$

and hence $\tau' = \tau$. $\square$

Definition 5.3. A relation R is a *bisimulation* between two labelled graphs $\mathbb{G} = (G, \twoheadrightarrow, \ell_0)$ and $\mathbb{H} = (H, \rightarrowtail, \ell_1)$ if $R \subseteq G \times H$ and the following conditions are satisfied (where aRb stands for $\langle a, b\rangle \in R$):

(1) For every $a \in G$ there is a $b \in H$ such that aRb.

(2) For every $b \in H$ there is a $a \in G$ such that aRb.
(3) Suppose that aRb. Then for every $x \in G$ such that $a \twoheadrightarrow x$ there is a $y \in H$ such that $b \rightarrowtail y$ and xRy.
(4) Suppose that aRb. Then for every $y \in H$ such that $b \rightarrowtail y$ there is an $x \in G$ such that $a \twoheadrightarrow x$ and xRy.
(5) If aRb then $\ell_0(a) = \ell_1(b)$.

Two labelled graphs are *bisimular* if there exists a bisimulation between them.

Theorem 5.4. (**CZFA**) *Let $\mathbb{G} = (G, \twoheadrightarrow, \ell_0)$ and $\mathbb{H} = (H, \rightarrowtail, \ell_1)$ be labelled graphs with labelled decorations d_0 and d_1, respectively.*
If $\mathbb{G}$ and $\mathbb{H}$ are bisimular then $d_0[G] = d_1[H]$.

Proof: Define a labelled graph $\mathbb{K} = (K, \rightarrow, \ell)$ by letting K be the set $\{\langle a, b\rangle : aRb\}$. For $\langle a, b\rangle, \langle a', b'\rangle \in K$ let $\langle a, b\rangle \rightarrow \langle a', b'\rangle$ iff $a \twoheadrightarrow a'$ or $b \rightarrowtail b'$, and put $\ell(\langle a, b\rangle) = \ell_0(a) = \ell_1(b)$. $\mathbb{K}$ has a unique labelled decoration d. Using a bisimulation R, one easily verifies that $d_0^*(\langle a, b\rangle) := d_0(a)$ and $d_1^*(\langle a, b\rangle) := d_1(b)$ are labelled decorations of $\mathbb{K}$ as well. Hence $d = d_0^* = d_1^*$, and thus $d_0[G] = d[K] = d_1[H]$. ☐

Corollary 5.5. (**CZFA**) *Two graphs are bisimular if and only if their decorations have the same image.*

Proof: One direction follows from the previous theorem. Now suppose we have graphs $\mathbb{G} = (G, \twoheadrightarrow)$ and $\mathbb{H} = (H, \rightarrowtail)$ with decorations d_0 and d_1, respectively, such that $d_0[G] = d_1[H]$. The define $R \subseteq G \times H$ by aRb iff $d_0(a) = d_1(b)$. One readily verifies that R is a bisimulation. ☐

Here is another useful fact:

Lemma 5.6. (**CZFA**) *If A is transitive set and $d : A \rightarrow V$ is a function such that $d(a) = \{d(x) : x \in a\}$ for all $a \in A$, then $d(a) = a$ for all $a \in A$.*

Proof: A can be considered the set of nodes of the graph $\mathbb{G}_A = (A, \rightarrowtail)$ where $a \rightarrowtail b$ iff $b \in a$ and $a, b \in A$. Since A is transitive, d is a decoration of $\mathbb{G}$. But so is the function $a \mapsto a$. Thus we get $d(a) = a$. ☐

5.2. **Systems.** In applications it is often useful to avail oneself of graphs that are classes rather than sets. By a *map* $\wp$ with domain M we mean a definable class function with domain M, and we will write $\wp : M \rightarrow V$.

Definition 5.7. A *labelled system* is a class M of nodes together with a labelling map $\wp : M \rightarrow V$ and a class E of edges consisting of ordered pairs of nodes. Furthermore, a system is required to satisfy that for each node $a \in M$, $\{b \in M : a \rightarrowtail b\}$ is a set, where $a \rightarrowtail b$ stands for $\langle a, b\rangle \in E$.

The labelled system is said to be Δ_0 if the relation between sets x and y defined by "$y = \{b \in M : a \rightarrowtail b$ for some $a \in x\}$" is Δ_0 definable.

We will abbreviate the labelled system by $\mathbb{M} = (M, \rightarrowtail, \wp)$.

Theorem 5.8. (**CZFA**+$\mathbf{IND}_\omega$) (Cf. [3], Theorem 1.10) *For every labelled system* $\mathbb{M} = (M, \rightarrowtail, \wp)$ *there exists a unique map* $d : M \to V$ *such that, for all* $a \in M$*:*

$$d(a) \;=\; \{d(b) \,:\, a \to b\} \,\cup\, \ell(a). \tag{2}$$

Proof: To each $a \in M$ we may associate a labelled graph $\mathbb{M}_a = (M_a, {}_a\!\rightarrowtail, \wp_a)$ with $M_a = \bigcup_{n\in\omega} X_n$, where $X_0 = \{a\}$ and $X_{n+1} = \{b : a \rightarrowtail b \text{ for some } a \in X_n\}$. The existence of the function $n \mapsto X_n$ is shown via recursion on ω, utilizing $\mathbf{IND}_\omega$ in combination with Strong Collection. The latter is needed to show that for every set Y, $\{b : a \rightarrowtail b \text{ for some } a \in Y\}$ is a set as well. And consequently to that M_a is a set. ${}_a\!\rightarrowtail$ is the restriction of $\rightarrowtail$ to nodes from M_a. That $E_a = \{\langle x, y\rangle \in M_a \times M_a : x \rightarrowtail y\}$ is a set requires Strong Collection, too. Further, let $\wp_a$ be the restriction of $\wp$ to M_a. Hence $\mathbb{M}_a$ is a set and we may apply Theorem 5.2 to conclude that $\mathbb{M}_a$ has a unique labelled decoration d_a. $d : M \to V$ is now obtained by patching together the function d_a with $a \in M$, that is $d = \bigcup_{a\in V} d_a$. One easily shows that two function d_a and d_b agree on $M_a \cap M_b$. For the uniqueness of d, notice that every other definable map d' satisfying (2) yields a function when restricted to M_a (Strong Collection) and thereby yields also a labelled decoration of $\mathbb{M}_a$; thus $d'(x) = \wp_a(x) = d(x)$ for all $x \in M_a$. And consequently to that, $d'(x) = d(x)$ for all $x \in M$. □

Corollary 5.9. (**CZFA**+Σ-$\mathbf{IND}_\omega$) *For every labelled system* $\mathbb{M} = (M, \rightarrowtail, \wp)$ *that is* Δ_0 *there exists a unique map* $d : M \to V$ *such that, for all* $a \in M$*:*

$$d(a) \;=\; \{d(b) \,:\, a \to b\} \,\cup\, \ell(a). \tag{3}$$

Proof: This follows by scrutinizing the proof of Theorem 5.8 and realizing that for a Δ_0 system one only needs Σ-$\mathbf{IND}_\omega$. □

Corollary 5.10. (**CZFA**) *Let* $\mathbb{M} = (M, \rightarrowtail, \wp)$ *be a labelled* Δ_0 *system such that for each* $a \in M$ *there is a function* $n \mapsto X_n$ *with domain* ω *such that* $X_0 = \{a\}$ *and* $X_{n+1} = \{b : a \rightarrowtail b$ *for some* $a \in X_n\}$. *Then there exists a unique map* $d : M \to V$ *such that, for all* $a \in M$*:*

$$d(a) \;=\; \{d(b) \,:\, a \to b\} \,\cup\, \ell(a). \tag{4}$$

Proof: In the proof of Theorem 5.8 we employed $\mathbf{IND}_\omega$ only once to ensure that $M_a = \bigcup_{n\in\omega} X_n$ is a set. This we get now for free from the assumptions. □

Theorem 5.11. (**CZFA** + $\mathbf{IND}_\omega$) (Cf. [3], Theorem 1.11) *Let* $\mathbb{M} = (M, \twoheadrightarrow, \wp)$ *be a labelled system whose sets of labels are subsets of the class* Y.

(1) If π *is a map with domain* Y *then there is a unique function* $\hat\pi$ *with domain* M *such that for each* $a \in M$

$$\hat\pi(a) \;=\; \{\hat\pi(b) \,:\, a \twoheadrightarrow b\} \;\cup\; \{\pi(x) \,:\, x \in \wp(a)\}.$$

(2) *Given a map $\hbar : Y \to M$, there is a unique map π with domain Y such that for all $y \in Y$,*

$$\pi(y) = \hat{\pi}(\hbar(y)).$$

Proof: For (1) let $\mathbb{M}_\pi = (M, \twoheadrightarrow, \wp_\pi)$ be obtained from $\mathbb{M}$ and $\pi : Y \to V$ by redefining the sets of labels so that for each node a

$$\wp_\pi(a) = \{\pi(x) : x \in \wp(a)\}.$$

Then the required unique map $\hat{\pi}$ is the unique labelled decoration of $\mathbb{M}_\pi$ provided by Theorem 5.8

For (2) let $\mathbb{M}^* = (M, \rightarrowtail)$ be the graph having the same nodes as $\mathbb{M}$, and all edges of $\mathbb{M}$ together with the edges $a \rightarrowtail \hbar(y)$ whenever $a \in M$ and $y \in \wp(a)$. By Theorem 5.8, $\mathbb{M}^*$ has a unique decoration map ρ. So for each $a \in M$

$$\rho(a) = \{\rho(b) : a \twoheadrightarrow b\} \cup \{\rho(\hbar(y)) : y \in \wp(a)\}.$$

Letting $\pi(y) := \rho(\hbar(y))$ for $y \in Y$, ρ is also a labelled decoration for the labelled system $\mathbb{M}_\pi$ so that $\rho = \hat{\pi}$ by (1), and hence $\pi(x) = \hat{\pi}(\hbar(x))$ for $x \in Y$. For the uniqueness of π let $\mu : M \to V$ satisfy $\mu(x) = \hat{\mu}(\hbar(x))$ for $x \in Y$. Then $\hat{\mu}$ is a decoration of $\mathbb{M}^*$ as well, so that $\hat{\mu} = \rho$. As a result $\mu(x) = \hat{\mu}(\hbar(x)) = \rho(\hbar(x)) = \pi(x)$ for $x \in Y$. Thus $\mu(x) = \pi(x)$ for all $x \in Y$. $\square$

Corollary 5.12. (**CZFA** + **Σ-IND$_\omega$**) *Let $\mathbb{M} = (M, \twoheadrightarrow, \wp)$ be a labelled system that is Δ_0 and whose sets of labels are subsets of the class Y.*

(1) *If π is a map with domain Y then there is a unique function $\hat{\pi}$ with domain M such that for each $a \in M$*

$$\hat{\pi}(a) = \{\hat{\pi}(b) : a \twoheadrightarrow b\} \cup \{\pi(x) : x \in \wp(a)\}.$$

(2) *Given a map $\hbar : Y \to M$ there is a unique map π with domain Y such that for all $x \in Y$,*

$$\pi(x) = \hat{\pi}(\hbar(x)).$$

Proof: The proof is the same as for Theorem 5.11, except that one utilizes Corollary 5.9 in place of Theorem 5.8. $\square$

Corollary 5.13. (**CZFA**) *Let $\mathbb{M} = (M, \twoheadrightarrow, \wp)$ be a labelled system that is Δ_0 and whose sets of labels are subsets of the class Y. Moreover suppose that for each $a \in M$ there is a function $n \mapsto X_n$ with domain ω such that $X_0 = \{a\}$ and $X_{n+1} = \{b : a \rightarrowtail b$ for some $a \in X_n\}$.*

(1) *If π is a map with domain Y then there is a unique function $\hat{\pi}$ with domain M such that for each $a \in M$*

$$\hat{\pi}(a) = \{\hat{\pi}(b) : a \twoheadrightarrow b\} \cup \{\pi(x) : x \in \wp(a)\}.$$

(2) *Given a map $\hbar : Y \to M$ there is a unique map π with domain Y such that for all $x \in Y$,*

$$\pi(x) = \hat{\pi}(\hbar(x)).$$

Proof: The proof is the same as for Theorem 5.11, except that one utilizes Corollary 5.10 in place of Theorem 5.8. □

5.3. **A Solution Lemma version of AFA.** **AFA** can be couched in more traditional mathematical terms. The labelled Anti-Foundation Axiom provides a nice tool for showing that systems of equations of a certain type have always unique solutions. In the terminology of [5] this is called the *solution lemma*. In [5], the Anti-Foundation Axiom is even expressed in terms of unique solutions to so-called *flat systems of equations*.

Definition 5.14. For a set Y let $\mathcal{P}(Y)$ be the class of subsets of Y. A triple $\mathcal{E} = (X, A, e)$ is said to be a *general flat system of equations* if X and A are any two sets, and $e : X \to \mathcal{P}(X \cup A)$, where the latter conveys that e is a function with domain X which maps into the class of all subsets of $X \cup A$. X will be called the set of *indeterminates* of $\mathcal{E}$, and A is called the set of *atoms* of $\mathcal{E}$. Let $e_v = e(v)$. For each $v \in X$, the set $b_v := e_v \cap X$ is called the set of indeterminates on which v immediately depends. Similarly, the set $c_v := e_v \cap A$ is called the set of atoms on which v immediately depends.

A *solution* to $\mathcal{E}$ is a function s with domain X satisfying

$$s_x = \{s_y : y \in b_x\} \cup c_x,$$

for each $x \in X$, where $s_x := s(x)$.

Theorem 5.15. (**CZFA**) *Every generalized flat system $\mathcal{E} = (X, A, e)$ has a unique solution.*

Proof: Define a labelled graph $\mathbb{H}$ by letting X be its set of nodes and its edges be of the form $x \twoheadrightarrow y$, where $y \in b_x$ for $x, y \in X$. Moreover, let $\ell(x) = c_x$ be the pertinent labelling function. By Theorem 5.2, $\mathbb{H}$ has a unique labelled decoration d. Then

$$d(x) = \{d(y) : y \in b_x\} \cup \ell(x) = \{d(y) : y \in b_x\} \cup c_x,$$

and thus d is a solution to $\mathcal{E}$. One easily verifies that every solution s to $\mathcal{E}$ gives rise to a decoration of $\mathbb{H}$. Thus there exists exactly one solution to $\mathcal{E}$. □

Because of the flatness condition, i.e. $e : X \to \mathcal{P}(X \cup A)$, the above form of the Solution Lemma is often awkward to use. A much more general form of it is proved in [5]. The framework in [5], however, includes other objects than sets, namely a proper class of urelements, whose raison d'etre is to serve as an endless supply of indeterminates on which one can perform the operation of substitution. Given a set X of urelements one defines the class of X-sets which are those sets that use only urelements from X in their build-up. For a function $f : X \to V$ on these indeterminates one can then define a substitution operation sub_f on the X-sets. For an X-set a, $sub_f(a)$ is obtained from a by substituting $f(x)$ for x everywhere in the build-up of a.

For want of urelements, the approach of [5] is not directly applicable in our set theories, though it is possible to model an extended universe of sets with a proper class of urelements within **CZFA**. This will require a class defined as the greatest fixed point of an operator, a topic I shall intersperse now.

5.4. **Greatest fixed points of operators.** The theory of greatest fixed points was initiated by Aczel in [3].

Definition 5.16. Let Φ be a class operator, i.e. $\Phi(X)$ is a class for each class X. Φ is *set continuous* if for each class X

$$\Phi(X) \;=\; \bigcup\{\Phi(x):\; x \textit{ is a set with } x \subseteq X\}. \tag{5}$$

Note that a set continuous operator is monotone, i.e., if $X \subseteq Y$ then $\Phi(X) \subseteq \Phi(Y)$.

In what follows, I shall convey that x is a set by $x{\in}V$. If Φ is a set continuous operator let

$$J_\Phi \;=\; \bigcup\{x{\in}V:\; x \subseteq \Phi(x)\}.$$

A set continuous operator Φ is Δ_0 if the relation "$y \in \Phi(x)$" between sets x and y is Δ_0 definable. Notice that J_Φ is a Σ_1 class if Φ is a Δ_0 operator.

Theorem 5.17. (**CZF**$^-$ + **RDC**) (Cf. [3], Theorem 6.5) *If Φ is a set continuous operator and $J = J_\Phi$ then*

(1) $J \subseteq \Phi(J)$,
(2) If $X \subseteq \Phi(X)$ then $X \subseteq J$,
(3) J is the largest fixed point of Φ.

Proof: (1): Let $a{\in}J$. Then $a{\in}x$ for for some set x such that $x \subseteq \Phi(x)$. It follows that $a \in \Phi(J)$ as $x \subseteq J$ and Φ is monotone.

(2): Let $X \subseteq \Phi(X)$ and let $a{\in}X$. We like to show that $a{\in}J$. We first show that for each set $x \subseteq X$ there is a set $c_x \subseteq X$ such that $x \subseteq \Phi(c_x)$. So let $x \subseteq X$. Then $x \subseteq \Phi(X)$ yielding

$$\forall y{\in}x\; \exists u\; [y{\in}\Phi(u) \;\wedge\; u \subseteq X].$$

By Strong Collection there is a set A such that

$$\forall y{\in}x\; \exists u{\in}A\; [y{\in}\Phi(u) \;\wedge\; u \subseteq X] \;\wedge\; \forall u{\in}A\; \exists y{\in}x\; [y{\in}\Phi(u) \;\wedge\; u \subseteq X].$$

Letting $c_x = \bigcup A$, we get $c_x \subseteq X \;\wedge\; x \subseteq \Phi(c_x)$ as required.

Next we use **RDC** to find an infinite sequence $x_0, x_1, \ldots$ of subsets of X such that $x_0 = \{a\}$ and $x_n \subseteq \Phi(x_{n+1})$. Let $x^* = \bigcup_n x_n$. Then x^* is a set and if $y{\in}x^*$ then $y{\in}x_n$ for some n so that $y \in x_n \subseteq \Phi(x_{n+1}) \subseteq \Phi(x^*)$. Hence $x^* \subseteq \Phi(x^*)$. As $a{\in}x_0 \subseteq x^*$ it follows that $a{\in}J$.

(3): By (1) and the monotonicity of Φ

$$\Phi(J) \subseteq \Phi(\Phi(J)).$$

Hence by (2) $\Phi(J) \subseteq J$. This and (1) imply that J is a fixed point of Φ. By (2) it must be the greatest fixed point of Φ. $\square$

If it exists and is a set, the largest fixed point of an operator Φ will be called the set *coinductively defined* by Φ.

Theorem 5.18. ($\mathbf{CZF}^- + \Delta_0$-$\mathbf{RDC}$) *If Φ is a set continuous Δ_0 operator and $J = J_\Phi$ then*

(1) $J \subseteq \Phi(J)$,
(2) If X is a Σ_1 class and $X \subseteq \Phi(X)$ then $X \subseteq J$,
(3) J is the largest Σ_1 fixed point of Φ.

Proof: This is the same proof as for Theorem 5.17, noticing that Δ_0-$\mathbf{RDC}$ suffices here. $\square$

In applications, set continuous operators Φ often satisfy an additional property. Φ will be called *fathomable* if there is a partial class function q such that whenever $a \in \Phi(x)$ for some set x then $q(a) \subseteq x$ and $a \in \Phi(q(a))$. For example, deterministic inductive definitions are given by fathomable operators.

If the graph of q is also Δ_0 definable we will say that Φ is a fathomable set continuous Δ_0 operator.

For fathomable operators one can dispense with $\mathbf{RDC}$ and Δ_0-$\mathbf{RDC}$ in Theorems 5.17 and 5.18 in favour of $\mathbf{IND}_\omega$ and Σ-$\mathbf{IND}_\omega$, respectively.

Corollary 5.19. ($\mathbf{CZF}^- + \mathbf{IND}_\omega$) *If Φ is a set continuous fathomable operator and $J = J_\Phi$ then*

(1) $J \subseteq \Phi(J)$,
(2) If $X \subseteq \Phi(X)$ then $X \subseteq J$,
(3) J is the largest fixed point of Φ.

Proof: In the proof of Theorem 5.17, $\mathbf{RDC}$ was used for (2) to show that for every class X with $X \subseteq \Phi(X)$ it holds $X \subseteq J$. Now, if $a \in X$, then $a \in \Phi(u)$ for some set $u \subseteq X$, as Φ is set continuous, and thus $a \in \Phi(q(a))$ and $q(a) \subseteq X$. Using $\mathbf{IND}_\omega$ and Strong Collection one defines a sequence $x_0, x_1, \ldots$ by $x_0 = \{a\}$ and $x_{n+1} = \bigcup\{q(v) : v \in x_n\}$. We use induction on ω to show $x_n \subseteq X$. Obviously $x_0 \subseteq X$. Suppose $x_n \subseteq X$. Then $x_n \subseteq \Phi(X)$. Thus for every $v \in x_n$, $q(v) \subseteq X$, and hence $x_{n+1} \subseteq X$. Let $x^* = \bigcup_n x_n$. Then $x^* \subseteq X$. Suppose $u \in x^*$. Then $u \in x_n$ for some n, and hence as $u \in \Phi(X)$, $u \in \Phi(q(u))$. Thus $q(u) \subseteq x_{n+1} \subseteq x^*$, and so $u \in \Phi(x^*)$. As a result, $a \in x^* \subseteq \Phi(x^*)$, and hence $a \in J$. $\square$

Corollary 5.20. ($\mathbf{CZF}^- + \Sigma$-$\mathbf{IND}_\omega$) *If Φ is a set continuous fathomable Δ_0 operator and $J = J_\Phi$ then*

(1) $J \subseteq \Phi(J)$,
(2) If X is Σ_1 and $X \subseteq \Phi(X)$ then $X \subseteq J$,
(3) J is the largest Σ_1 fixed point of Φ.

Proof: If the graph of q is Δ_0 definable, Σ-$\mathbf{IND}_\omega$ is sufficient to define the sequence $x_0, x_1, \ldots$. $\square$

For special operators it is also possible to forgo Σ-$\mathbf{IND}_\omega$ in favour of TRANS.

Corollary 5.21. ($\mathbf{CZF}^-$ + TRANS) *Let Φ be a set continuous fathomable Δ_0 operator such that q is a total map and $q(a) \subseteq \mathbf{TC}(\{a\})$ for all sets a. Let $J = J_\Phi$. Then*

(1) $J \subseteq \Phi(J)$,
(2) If X *is* Δ_0 *and* $X \subseteq \Phi(X)$ *then* $X \subseteq J$,
(3) J *is the largest* Δ_0 *fixed point of* Φ.

Proof: (1) is proved as in Theorem 5.17. For (2), suppose that X is a class with $X \subseteq \Phi(X)$. Let $a \in X$. Define a sequence of sets $x_0, x_1, \ldots$ by $x_0 = \{a\}$ and $x_{n+1} = \bigcup\{q(v) : v{\in}x_n\}$ as in Corollary 5.19. But without Σ-$\mathbf{IND}_\omega$, how can we ensure that the function $n \mapsto x_n$ exists? This can be seen as follows. Define

$$\begin{aligned} D_n &= \{f \in {}^{n+1}\mathbf{TC}(\{a\}) : f(0) = a \wedge \forall i \in n\,[f(i+1) \in q(f(i))]\}, \\ E_n &= \{f(n) : f \in D_n\}. \end{aligned}$$

The function $n \mapsto E_n$ exists by Strong Collection. Moreover, $E_0 = \{a\}$ and $E_{n+1} = \bigcup\{q(v) : v \in E_n\}$ as can be easily shown by induction on n; thus $x_n = E_n$. The remainder of the proof is as in Corollary 5.19.

For (3), note that $J = \{a : a \in \Phi(q(a))\}$ and thus J is Δ_0. □

Remark 5.22. It is an open problem whether the above applications of the dependent choices axiom are necessary for the general theory of greatest fixed points.

5.5. **Generalized systems of equations in an expanded universe.** Before we can state the notion of a general systems of equations we will have to emulate urelements and the sets built out of them in the set theory **CZFA** with pure sets. To this end we employ the machinery of greatest fixed points of the previous subsection. We will take the sets of the form $\langle 1, x\rangle$ to be the urelements and call them $*$-*urelements*. The class of $*$-urelements will be denoted by $\mathcal{U}$. Certain sets built from them will be called the $*$-*sets*. If $a = \langle 2, u\rangle$ let $a^* = u$. The elements of a^* will be called the $*$-elements of a. Let the $*$-sets be the largest class of sets of the form $a = \langle 2, u\rangle$ such that each $*$-element of a is either a $*$-urelement or else a $*$-set. To bring this under the heading of the previous subsection, define

$$\Phi^*(X) \;=\; \{\langle 2, u\rangle : \forall x{\in}u\,\big[(x{\in}X \wedge x \in \mathrm{TWO}) \vee x \text{ is a } *\text{-urelement}\big]\},$$

where TWO is the class of all ordered pairs of the form $\langle 2, v\rangle$. Obviously, Φ^* is a set-continuous operator. That Φ^* is fathomable can be seen by letting

$$q(a) \;=\; \{v \in a^* : v \in \mathrm{TWO}\}.$$

Notice also that Φ^* has a Δ_0 definition.

The $*$-sets are precisely the elements of J_{Φ^*}. Given a class of Z of $*$-urelements we will also define the class of Z-*sets* to be the largest class of $*$-sets such that every $*$-urelement in a Z-set is in Z. We will use the notation $V[Z]$ for the class of Z-sets.

Definition 5.23. A *general system of equations* is a pair $\mathcal{E} = (X, e)$ consisting of a set $X \subseteq \mathcal{U}$ (of indeterminates) and a function

$$e : X \to V[X].$$

The point of requiring e to take values in $V[X]$ is that thereby e is barred from taking $*$-urelements as values and that all the values of e are sets which use only $*$-urelements from X in their build-up. In consequence, one can define a substitution operation on the values of e.

Theorem 5.24. (**CZFA**) (Substitution Lemma) *Let Y be a Δ_0 class such that $Y \subseteq \mathcal{U}$. For each map $\rho : Y \to V$ there exists a unique operation sub_ρ that assigns to each $a \in V[Y]$ a set $sub_\rho(a)$ such that*

$$sub_\rho(a) \;=\; \{sub_\rho(x) \,:\, x \in a^* \cap V[Y]\} \,\cup\, \{\rho(x) \,:\, x \in a^* \cap Y\}. \tag{6}$$

Proof: The class $V[Y]$ forms the nodes of a labelled Δ_0 system $\mathbb{M}$ with edges $a \rightarrowtail b$ for $a, b \in V[Y]$ whenever $b \in a^*$, and labelling map $\wp(a) = a^* \cap Y$. By Corollary 5.13 there exists a unique map $\hat{\rho} : V[Y] \to V$ such that for each $a \in V[Y]$,

$$\hat{\rho}(a) \;=\; \{\hat{\rho}(x) : x \in a^* \cap V[Y]\} \,\cup\, \{\rho(x) : x \in a^* \cap Y\}. \tag{7}$$

Put $sub_\rho(a) := \hat{\rho}(a)$. Then sub_ρ satisfies (6). Since the equation (7) uniquely determines $\hat{\rho}$ it follows that sub_ρ is uniquely determined as well. □

Definition 5.25. Let $\mathcal{E}$ be a general system of equations as in Definition 5.23. A *solution* to $\mathcal{E}$ is a function $s : X \to V$ satisfying, for all $x \in X$,

$$s(x) \;=\; sub_s(\,e_x), \tag{8}$$

where $e_x := e(x)$.

Theorem 5.26. (**CZFA**) (Solution Lemma) *Let $\mathcal{E}$ be a general system of equations as in Definition 5.23. Then $\mathcal{E}$ has a unique solution.*

Proof: The class $V[X]$ provides the nodes for a labelled Δ_0 system $\mathbb{M}$ with edges $b \rightarrowtail c$ for $b, c \in V[X]$ whenever $c \in b^*$, and with a labelling map $\wp(b) = b^* \cap X$. Since $e : X \to V[X]$, we may employ Corollary 5.13 (with $Y = X$). Thus there is a unique function π and a unique function $\hat{\pi}$ such that

$$\pi(x) \;=\; \hat{\pi}(\,e_x) \tag{9}$$

for all $x \in X$, and

$$\hat{\pi}(a) \;=\; \{\hat{\pi}(b) :: b \in a^*\} \,\cup\, \{\pi(x) \,:\, x \in a^* \cap X\}. \tag{10}$$

In view of Theorem 5.24, we get $\hat{\pi} = sub_\pi$ from (10). Thus letting $s := \pi$, (9) then yields the desired equation $s(x) = sub_s(\,e_x)$ for all $x \in X$. Further, s is unique owing to the uniqueness of π in (9). □

Remark 5.27. The framework in which **AFA** is studied in [5] is a set theory with a proper class of urelements $\mathcal{U}$ that also features an *axiom of plenitude* which is the conjunction of the following sentences:

$$\begin{aligned}
&\forall a \forall b\, \mathsf{new}(a,b) \in \mathcal{U},\\
&\forall a \forall a' \forall b \forall b' \left[\mathsf{new}(a,b) = \mathsf{new}(a',b') \to a = a' \,\wedge\, b = b'\right],\\
&\forall a \forall b \left[b \subseteq \mathcal{U} \to \mathsf{new}(a,b) \notin b\right],
\end{aligned}$$

where new is a binary function symbol. It is natural to ask whether a version of **CZFA** with urelements and an axiom of plenitude would yield any extra strength. That such a theory is not stronger than **CZFA** can be easily seen by modelling the urelements and sets of [5] inside **CZFA** by the $*$-urelements and the $*$-sets, respectively. To interpret the function symbol new define

$$\mathsf{new}^*(a,b) := \langle 1, \langle a, \langle b, b^r \rangle\rangle\rangle,$$

where $b^r = \{r \in \mathbf{TC}(b) : r \notin r\}$. Obviously, $\mathsf{new}^*(a,b)$ is a $*$-urelement and new^* is injective. Moreover, $\mathsf{new}^*(a,b) \in b$ would imply $\mathsf{new}^*(a,b) \in \mathbf{TC}(b)$ and thus $b^r \in \mathbf{TC}(b)$. The latter yields the contradiction $b^r \notin b^r \wedge b^r \in b^r$. As a result, $\mathsf{new}^*(a,b) \notin b$. Interpreting new by new^* thus validates the axiom of plenitude, too.

5.6. **Streams, coinduction, and corecursion.** In the following I shall demonstrate the important methods of coinduction and corecursion in a setting which is not too complicated but still demonstrates the general case in a nutshell. The presentation closely follows [5].

Let A be some set. By a *stream* over A we mean an ordered pair $s = \langle a, s' \rangle$ where $a \in A$ and s' is another stream. We think of a stream as being an element of A followed by another stream. Two important operations performed on streams s are taking the first element $1^{st}(s)$ which gives an element of A, and taking its second element $2^{nd}(s)$, which yields another stream. If we let A^∞ be the streams over A, then we would like to have

$$A^\infty = A \times A^\infty. \tag{11}$$

In set theory with the foundation axiom, equation (11) has only the solution $A = \emptyset$. With **AFA**, however, not only can one show that (11) has a solution different from $\emptyset$ but also that it has a largest solution, the latter being the largest fixed point of the operator $\Gamma_A(Z) = A \times Z$. This largest solution to Γ_A will be taken to be the set of streams over A and be denoted by A^∞, thus rendering A^∞ a *coinductive* set. Moreover, it will be shown that A^∞ possesses a "recursive" character despite the fact that there is no "base case". For instance, it will turn out that one can define a function

$$zip : A^\infty \times A^\infty \to A^\infty$$

such that for all $s, t \in A^\infty$

$$zip(s,t) = \langle 1st(s), \langle 1^{st}(t), zip(2^{nd}(s), 2^{nd}(t))\rangle\rangle. \tag{12}$$

As its name suggests, *zip* acts like a zipper on two streams. The definition of *zip* in (12) is an example for definition by *corecursion* over a coinductive set.

Theorem 5.28. (**CZFA**) *For every set A there is a largest set Z such that $Z \subseteq A \times Z$. Moreover, Z satisfies $Z = A \times Z$, and if A is inhabited then so is Z.*

Proof: Let F be the set of functions from $\mathbb{N} := \omega$ to A. For each such f, we define another function $f^+ : \mathbb{N} \to \mathbb{N}$ by

$$f^+(n) = f(n+1).$$

For each $f \in F$ let x_f be an indeterminate. We would like to solve the system of equations given by

$$x_f \;=\; \langle f(0), x_{f^+}\rangle.$$

Solving these equations is equivalent to solving the equations

$$\begin{aligned} x_f &= \{y_f, z_f\}; \\ y_f &= \{f(0)\} \\ z_f &= \{f(0), x_{f^+}\}, \end{aligned} \tag{13}$$

where y_f and z_f are further indeterminates. Note that $f(0)$ is an element of A. To be precise, let $x_f = \langle 0, f\rangle$, $y_f = \langle 1, f\rangle$, and $z_f = \langle 2, f\rangle$. Solving (13) amounts to the same as finding a labelled decoration for the labelled graph

$$\mathbb{S}_A = (S, \twoheadrightarrow, \ell)$$

whose set of nodes is

$$S = \{x_f \,:\, f \in F\} \cup \{y_f \,:\, f \in F\} \cup \{z_f \,:\, f \in F\}$$

and whose edges are given by $x_f \twoheadrightarrow y_f$, $x_f \twoheadrightarrow z_f$, $z_f \twoheadrightarrow x_{f^+}$. Moreover, the labelling function ℓ is defined by $\ell(x_f) = \emptyset$, $\ell(y_f) = \{f(0)\}$, $\ell(z_f) = \{f(0)\}$ for all $f \in F$. By the labelled Anti-Foundation Axiom, Theorem 5.2, $\mathbb{S}_A$ has a labelled decoration d and we thus get

$$d(x_f) \;=\; \langle f(0), d(x_{f^+})\rangle. \tag{14}$$

Let $A^\infty = \{d(x_f) \,:\, f \in F\}$. By (14), we have $A^\infty \subseteq A \times A^\infty$. Thus A^∞ solves the equation $Z \subseteq A \times Z$.

To check that $A \times A^\infty \subseteq A^\infty$ holds also, let $a \in A$ and $t \in A^\infty$. By the definition of A^∞, $t = d(x_f)$ for some $f \in F$. Let $g : \mathbb{N} \to A$ be defined by $g(0) = a$ and $g(n+1) = f(n)$. Then $g^+ = f$, and thus $d(x_g) = \langle a, d(x_f)\rangle = \langle a, t\rangle$, so $\langle a, t\rangle \in A^\infty$.

If A contains an element a, then $f_a \in F$, where $f_a : \mathbb{N} \to A$ is defined by $f_a(n) = a$. Hence $d(x_{f_a}) \in A^\infty$, so A^∞ is inhabited, too.

Finally it remains to show that A^∞ is the largest set Z satisfying $Z \subseteq A \times Z$. So suppose that W is a set so that $W \subseteq A \times W$. Let $v \in W$. Define $f_v : \mathbb{N} \to A$ by

$$f_v(n) \;=\; 1^{st}(sec^n(v)),$$

where $sec^0(v) = v$ and $sec^{n+1}(v) = 2^{nd}(sec^n(v))$. Then $f_v \in F$, and so $d(x_{f_v}) \in A^\infty$. We claim that for all $v \in W$, $d(x_{f_v}) = v$. Notice first that for $w = 2^{nd}(v)$, we have $sec^n(w) = sec^{n+1}(v)$ for all $n \in \mathbb{N}$, and thus $f_w = (f_v)^+$. It follows that

$$\begin{aligned} d(x_{f_v}) &= \langle 1^{st}(v), d(x_{(f_v)^+})\rangle \\ &= \langle 1^{st}(v), d(x_{f_w})\rangle \\ &= \langle 1^{st}(v), d(x_{f_{2^{nd}(v)}})\rangle. \end{aligned} \tag{15}$$

W gives rise to a labelled subgraph $\mathbb{T}$ of $\mathbb{S}$ whose set of nodes is

$$T := \{x_{f_v} \,:\, v \in W\} \cup \{y_{f_v} \,:\, v \in W\} \cup \{z_{f_v} \,:\, v \in W\},$$

and wherein the edges and the labelling function are obtained from $\mathbb{S}$ by restriction to nodes from T. The function d' with $d'(x_{f_v}) = v$, $d'(y_{f_v}) = \{1^{st}(v)\}$, and $d'(z_{f_v}) = \{1^{st}(v), 2^{nd}(v)\}$ is obviously a labelled decoration of $\mathbb{T}$. By (15), d restricted to T is a labelled decoration of $\mathbb{T}$ as well. So by Theorem 5.2, $v = d'(x_{f_v}) = d(x_{f_v})$ for all $v \in W$, and thus $W \subseteq A^\infty$. □

Remark 5.29. Rather than applying the labelled Anti-Foundation Axiom one can utilize the solution lemma for general systems of equations (Theorem 5.26) in the above proof of Theorem 5.28. To this end let $B = \mathbf{TC}(A)$, $x_f = \langle 1, \langle 0, f\rangle\rangle$ for $f \in F$ and $x_b = \langle 1, \langle 1, b\rangle\rangle$ for $b \in B$. Set $X := \{x_f : f \in F\} \cup \{x_b : b \in B\}$. Then $X \subseteq \mathcal{U}$ and $\{x_f : f \in F\} \cap \{x_b : b \in B\} = \emptyset$.

Next define the unordered $*$-pair by $\{c, d\}^* = \langle 2, \{c, d\}\rangle$ and the ordered $*$-pair by $\langle c, d\rangle^* = \{\{c\}^*, \{c, d\}^*\}^*$. Note that with $c, d \in V[X]$ one also has $\{c, d\}^*, \langle c, d\rangle^* \in V[X]$.

Let $\mathcal{E} = (X, e)$ be the general system of equations with $e(x_f) = \langle x_{f(0)}, x_{f^+}\rangle^*$ for $f \in F$ and $e(x_b) = \langle 2, \{x_u \ \ u \in b\}\rangle$ for $b \in B$. Then $e : X \to V[X]$. By Theorem 5.26 there is a unique function $s : X \to V$ such that

$$s(x_b) = sub_s(e(x_b)) = \{s(x_u) : u \in b\} \quad \text{for } b \in B, \tag{16}$$

$$s(x_f) = sub_s(e(x_f)) = \langle s(x_{f(0)}), s(x_{f^+})\rangle \quad \text{for } f \in F. \tag{17}$$

From (16) and Lemma 5.6 it follows $s(x_b) = b$ for all $b \in B$, and thus from (17) it ensues that $s(x_f) = \langle f(0), s(x_{f^+})\rangle$ for $f \in F$. From here on one can proceed further just as in the proof of Theorem 5.26.

As a corollary to Theorem 5.28 one gets the following *coinduction principle* for A^∞.

Corollary 5.30. (**CZFA**) *If a set Z satisfies $Z \subseteq A \times Z$, then $Z \subseteq A^\infty$.*

Proof: This follows from the fact that A^∞ is the largest such set. □

The pivotal property of inductively defined sets is that one can define functions on them by structural recursion. For coinductively defined sets one has a dual principle, *corecursion*, which allows one to define functions mapping into the coinductive set.

Theorem 5.31. (**CZFA**) (Corecursion Pinciple for Streams). *Let C be an arbitrary set. Given functions $g : C \to A$ and $h : C \to C$ there is a unique function $f : C \to A^\infty$ satisfying*

$$f(c) = \langle g(c), f(h(c))\rangle \tag{18}$$

for all $c \in C$.

Proof: For each $c \in C$ let x_c, y_c, z_c be different indeterminates. To be precise, let $x_c = \langle 0, c\rangle$, $y_c = \langle 1, c\rangle$, and $z_c = \langle 2, c\rangle$ for $c \in C$. This time we would like to solve the system of equations given by

$$x_c = \langle g(c), x_{h(c)}\rangle.$$

Solving these equations is equivalent to solving the equations

$$\begin{array}{rcl} x_c &=& \{y_c, z_c\}; \\ y_c &=& \{g(c)\} \\ z_c &=& \{g(c), x_{h(c)}\}. \end{array} \tag{19}$$

Solving (19) amounts to the same as finding a labelled decoration for the labelled graph

$$\mathbb{S}_C = (S_C, \twoheadrightarrow, \ell_C)$$

whose set of nodes is

$$S_C = \{x_c : c \in C\} \cup \{y_c : c \in C\} \cup \{z_c : c \in C\}$$

and whose edges are given by $x_c \twoheadrightarrow y_c$, $x_c \twoheadrightarrow z_c$, $z_c \twoheadrightarrow x_{h(c)}$. Moreover, the labelling function ℓ_C is defined by $\ell_C(x_b) = \emptyset$, $\ell_C(y_b) = \{g(b)\}$, $\ell_C(z_b) = \{g(b)\}$ for all $b \in C$. By the labelled Anti-Foundation Axiom, Theorem 5.2, $\mathbb{S}_C$ has a labelled decoration $\jmath$ and we thus get

$$\jmath(x_c) \;=\; \langle g(c), \jmath(x_{h(c)})\rangle. \tag{20}$$

Letting the function f with domain C be defined by $f(c) := \jmath(x_c)$, we get from (20) that

$$f(c) \;=\; \langle g(c), f(h(c))\rangle \tag{21}$$

holds for all $c \in C$. As $\mathbf{ran}(f) \subseteq A \times \mathbf{ran}(f)$, Corollary 5.30 yields $\mathbf{ran}(f) \subseteq A^\infty$, thus $f : C \to A^\infty$.

It remains to show that f is uniquely determined by (21). So suppose $f' : C \to A^\infty$ is another function satisfying $f'(c) = \langle g(c), f'(h(c))\rangle$ for all $c \in C$. Then the function $\jmath'$ with $\jmath'(x_c) = f'(c)$, $\jmath'(y_c) = \{g(c)\}$, and $\jmath'(z_c) = \{g(c), f'(h(c))\}$ would give another labelled decoration of $\mathbb{S}_C$, hence $f(c) = \jmath(x_c) = \jmath'(x_c) = f'(x_c)$, yielding $f = f'$. $\square$

Example 1. Let $k : A \to A$ be arbitrary. Then k gives rise to a unique function $map_k : A^\infty \to A^\infty$ satisfying

$$map_k(s) \;=\; \langle k(1^{st}(s)), map_k(2^{nd}(s))\rangle. \tag{22}$$

For example, if $A = \mathbb{N}$, $k(n) = 2n$, and $s = \langle 3, \langle 6, \langle 9, \ldots\rangle\rangle\rangle$, then $map_k(s) = \langle 6, \langle 12, \langle 18, \ldots\rangle\rangle\rangle$. To see that map_k exists, let $C = A^\infty$ in Theorem 5.31, $g : A^\infty \to A$ be defined by $g(s) = k(1^{st}(s))$, and $h : A^\infty \to A^\infty$ be the function $h(s) = 2^{nd}(s)$. Then map_k is the unique function f provided by Theorem 5.31.

Example 2. Let $\nu : A \to A$. We want to define a function

$$iter_\nu : A \to A^\infty$$

which "iterates" ν such that $iter_\nu(a) = \langle a, iter_\nu(\nu(a))\rangle$ for all $a \in A$. If, for example $A = \mathbb{N}$ and $\nu(n) = 2n$, then $iter_\nu(7) = \langle 7, \langle 14, \langle 28, \ldots\rangle\rangle\rangle$. To arrive at $iter_\nu$ we employ Theorem 5.31 with $C = A^\infty$, $g : C \to A$, and $h : C \to C$, where $g(s) = \nu(1^{st}(s))$ and $h = map_\nu$, respectively.

Outlook. It would be desirable to develop the theory of corecursion of [5] (in particular Theorem 17.5) and the final coalgebra theorem of [3] in full generality within **CZFA** and extensions. It appears that the first challenge here is to formalize parts of category theory in constructive set theory. Due to page restrictions this cannot not be done in the present paper.

References

[1] P. Aczel: *The Type Theoretic Interpretation of Constructive Set Theory*, in: MacIntyre, A. Pacholski, L., Paris, J. (eds.), *Logic Colloquium '77* (North–Holland, Amsterdam, 1978) 55–66.

[2] P. Aczel: *The Type Theoretic Interpretation of Constructive Set Theory: Choice Principles*, in: Troelstra, A. S., van Dalen, D. (eds.), *The L.E.J. Brouwer Centenary Symposium*, (North–Holland, Amsterdam, 1982) 1–40.

[3] P. Aczel: *Non-well-founded sets.* CSLI Lecture Notes 14 (CSLI Publications, Stanford, 1988).

[4] J. Barwise: *Admissible sets and structures. An approach to definability theory*, (Springer Verlag, Berlin, 1975).

[5] J. Barwise, L. Moss: *Vicious circles.* CSLI Lecture Notes 60 (CSLI Publications, Stanford, 1996).

[6] M. Beeson: *Foundations of Constructive Mathematics*, (Springer Verlag, Berlin 1985).

[7] A. Cantini: *On the relation between choice and comprehension principles in second order arithmetic.* Journal of Symbolic Logic, vol. 51 (1986) 360-373.

[8] L. Crosilla: *Realizability interpretations for subsystems of* **CZF** *and proof theoretic strength* (Technical Report, School of Mathematics, University of Leeds, 1998).

[9] L. Crosilla: *Realizability models for constructive set theories wih restricted induction,* PhD Thesis (School of Mathematics, University of Leeds, 2000).

[10] L. Crosilla and M. Rathjen: *Inaccessible set axioms may have little consistency strength* Annals of Pure and Applied Logic 115 (2002) 33–70.

[11] S. Feferman: *Systems of predicative analysis*, Journal of Symbolic Logic 29 (1964) 1–30.

[12] S. Feferman: *Systems of predicative analysis II. Representations of ordinals*, Journal of Symbolic Logic 33 (1968) 193–220.

[13] S. Feferman: *A language and axioms for explicit mathematics.* In: J.N. Crossley (ed.): *Algebra and Logic*, Lecture Notes in Math. 450 (Springer, Berlin 1975) 87–139.

[14] S. Feferman: *Theories of finite type related to mathematical practice.* In: J. Barwise (ed.): *Handbook of Mathematical Logic* (North Holland, Amsterdam, 1977) 913–971.

[15] L. Hallnäs: *Non wellfounded sets: limits of wellfounded approximations.* Uppsala University, Department of Mathematics, Report No. 12 (1986) 16 pages.

[16] M. Forti, F. Honsell: *Set theory with free construction principles.* Annali Scuola Normale Supeiore di Pisa, Classe di Scienze 10 (1983) 493-522.

[17] G. Jäger: *Fixed points in Peano arithmetic with ordinals.* Annals of Pure and Applied Logic 60 (1993) 119–132.

[18] G. Kreisel: *Ordinal logics and the characterization of informal concepts of proof.* In: *Proceedings of the 1958 International Congress of Mathematicians*, (Edinburgh,1960) 289–299.

[19] S. Kripke: *Outline of a Theory of Truth.* The Journal of Philosophy 72 (1975) 53–81.

[20] D. Lewis: *Convention, A Philosophical Study* (Harvard University Press, Cambridge, MA, 1969).

[21] J. Lindström: *A construction of non-well-founded sets within Martin-Löf type theory.* Journal of Symbolic Logic 54 (1989) 57–64.

[22] P. Martin-Löf: *Intuitionistic Type Theory*, (Bibliopolis, Naples, 1984).

[23] J. Myhill: *Constructive Set Theory*, Journal of Symbolic Logic 40 (1975) 347–382.

[24] M. Rathjen: *A proof–theoretic characterization of the primitive recursive set functions.* Journal of Symbolic Logic 57 (1992) 954–969.

[25] M. Rathjen: *The strength of Martin-Löf type theory with a superuniverse. Part I*, Archive for Mathematical Logic 39 (2000) 1–39.
[26] M. Rathjen: *Kripke-Platek set theory and the anti-foundation axiom.* Mathematical Logic Quarterly 47 (2001) 435–440.
[27] M. Rathjen: *The anti-foundation axiom in constructive set theories.* In: G. Mints, R. Muskens (eds.): *Games, Logic, and Constructive Sets.* (CSLI Publications, Stanford, 2003) 87–108.
[28] B. Russell: *Principles of Mathematics.* (Cambridge University Press, Cambridge, 1903).
[29] B. Russell: *Mathematical logic as based on the theory of types.* American Journal of Mathematics 30 (1908) 222–262.
[30] K. Schütte: *Eine Grenze für die Beweisbarkeit der transfiniten Induktion in der verzweigten Typenlogik*, Archiv für Mathematische Logik und Grundlagenforschung 67 (1964) 45–60.
[31] K. Schütte: *Predicative well-orderings*, in: Crossley, Dummett (eds.), Formal systems and recursive functions (North Holland, Amsterdam, 1965) 176–184.
[32] K. Schütte: *Proof theory* (Springer, Berlin, 1977).
[33] S. Simpson: *Subsystems of second order arithmetic* (Springer, Berlin, 1999).
[34] H. Weyl: *Philosophy of Mathematics and Natural Sciences.* (Princeton University Press, Princeton, 1949).

School of Mathematics, University of Leeds, Leeds LS2 9JT, UK
E-mail address: rathjen@maths.leeds.ac.uk