Contemporary Mathematics

Permutations and Wellfoundedness: the True Meaning of the Bizarre Arithmetic of Quine's NF

Thomas Forster

ABSTRACT

It is shown that, according to NF, many of the assertions of ordinal arithmetic involving the T-function which is peculiar to NF turn out to be equivalent to the truth-in-certain-permutation-models of assertions which have perfectly sensible ZF-style meanings, such as: the existence of wellfounded sets of great size or rank, or the nonexistence of small counterexamples to the wellfoundedness of $\in$. Everything here holds also for NFU if the permutations are taken to fix all *urelemente*.

NF is Quine's system of set theory, axiomatised by extensionality and those instances of the naïve set existence scheme that are stratified. A formula of the language of set theory is stratified if the variables within it can be labelled with integers in such a way that all occurrences of each variable receive the same label, and if $x \in y$ appears in it then the label of x must be one less than the label of y, and if $x = y$ occurs then the label of x must be the same as the label of y.

In NF we can implement ordinals as isomorphism classes of wellorderings, and this has the effect that NO—the collection of all ordinals—is a set. If $\langle X, R\rangle$ is a wellordering of length α then $\{\{x\} : x \in X\}$ wellordered by the relation induced on it by R is likewise a wellordering (as one would expect) but of a length possibly not equal to α. Its length is said to be $T\alpha$. T is a definable endomorphism of the ordinals. It then turns out (using the definition of ordered pair appropriate to NF) that $\{\beta \in NO : \beta < \alpha\}$ is naturally of length $T^2\alpha$ rather than α. The expression '$\alpha = T\beta$' is stratified but *inhomogeneous*: that is to say 'α' is not of the same type as 'β'. Since the formula '$\alpha = T\alpha$' is not stratified, the graph of T is not *prima facie* a set, and we cannot use induction to prove $(\forall\alpha)(\alpha = T\alpha)$.

If the singleton function restricted to NO were a set then the graph of T would be a set and we would be able to prove by transfinite induction that T was the identity relation, and we would be able to obtain the Burali-Forti paradox. If the singleton function restricted to x exists we say x is **strongly cantorian**. If there is a bijection between x and $\{\{z\} : z \in x\}$ (not guaranteed to be the restriction of the singleton function) then we say x is **cantorian**.

We have just seen how the sethood of a particular collection defined by an unstratified condition (to wit $\{\alpha \in NO : \alpha = T\alpha\}$) would enable us to prove an assertion in ordinal-arithmetic-with-T (namely the Burali-Forti paradox). Are there converses? That is to say, do particular assertions about ordinal-arithmetic-with-T ever imply particular set-existence assertions? We have some familiarity with the phenomenon of assertions about ordinal-arithmetic-with-T if not actually *implying* nevertheless at least *having the same consistency strength as* particular formulæ that are—if not actually set existence assertions, at least combinatorial assertions about sets.

The results about consistency strength arise from considerations of Rieger-Bernays permutation models. If $\mathfrak{M} = \langle M, \in, = \rangle$ is a structure for the language of set theory, and σ is a permutation of M, then $\mathfrak{M}^\sigma = \langle M, \in_\sigma, = \rangle$ is another such structure, where $\in_\sigma$ is $\{\langle x, y\rangle : x \in \sigma(y)\}$. We can now write $\Diamond\phi$ for $(\exists\sigma)(V^\sigma \models \phi)$. (Notice that this means that σ must be a set of the model in question; see chapter 3 of [**2**] for details.) We use $\Box$ in the obvious dual sense. If $\phi \longleftrightarrow \Diamond\phi$ we say ϕ is **invariant**. All stratified sentences are invariant, though the converse is not true.

We have known for some time a number of equivalences of the form

$$\psi \longleftrightarrow \Diamond\phi$$

where ψ is an assertion in arithmetic-with-T and ϕ is an unstratified combinatorial assertion about $\in$. The axiom of counting (which says that $n = Tn$ for all finite ordinals n) affords an early and easy example, being equivalent to $\Diamond$($\mathbb{N}$ is strongly cantorian). In [**2**] I wrote up the nicest example then known to me of this phenomenon: theorem 3.1.28 on page 113 states that $(\forall n \in \mathbb{N})(n \leq Tn)$ (hereafter written "AxCount$_\leq$" as usual) is equivalent to the assertion that there is a permutation model in which $\{V_n : n < \omega\}$ is a set. That is to say, an assertion in ordinal-arithmetic-with-T was equivalent to the existence of a permutation model in which not only some cute combinatorial fact holds, but where that cute combinatorial fact is the existence of a particular definable wellfounded set whose existence makes perfect sense and has a standard meaning in a ZF context.

There is another class of unstratified combinatorial assertion we might be interested in. If $x = \{x\}$ we say x is a *Quine atom*. Quine atoms seem to be objects one could well do without: they represent failure of the axiom of foundation on a scale that seems quite uncalled-for. In [**6**] Dana Scott showed *inter alia* that it was consistent with NF that there should be no Quine atoms.

So we are going to try to tie together

(1) Good behaviour of the T-function;
(2) Existence of large wellfounded sets, or of wellfounded sets of high rank;
(3) Wellfoundedness of $\in$ restricted to small sets.

Let us discuss each of these briefly.

Good behaviour of T. The obvious way to formulate conjectures saying that T is well-behaved is to take a formula of ordinal arithmetic which one expects to be true, and mutilate it by prefixing some occurrences of some (ordinal) variables in it by the letter 'T'. The result of thus spraying an innocent formula with mutagens will be a (hopefully consistent) consequence of the (false) assumption that T is the identity. (For example the formula $(\forall\alpha)(\alpha \leq \alpha)$ gives rise to two formulæ: $(\forall\alpha)(T\alpha \leq \alpha)$ (which is sensible) and $(\forall\alpha)(\alpha \leq T\alpha)$ (which isn't)). Adopting the mutated formula as an axiom is thus an attempt to get T to behave in at least some ways as if it were the identity. We then sift through the results of these mutations for sensible conjectures. Those that are not obviously inconsistent represent more-or-less plausible attempts to recover as much normal behaviour of T as possible, and thus to get the ordinals of NF to behave as much like the ordinals of ZF as the hovering paradoxes will permit. This ought to be a sensible source of axioms, and it was seen as such from the first days of work on NF. The first axiom of this kind was Rosser's "axiom of counting" ("AxCount" for short) that $n = Tn$ for all natural numbers n (which implies the consistency of NF). The second was Orey's axiom that if $\alpha = T\alpha$, then $\beta = T\beta$ for all smaller β. It implies the first, and a lot more besides.

Some mutated assertions that arise in this way make sensible conjectures which invite investigation on their own merits. However, even for those that—like '$(\forall\alpha)(\alpha \leq T\alpha)$'—are obviously false that is not necessarily the end of the story. There is a systematic way of weakening these extreme assertions into sensible conjectures. This first arose in the context of Boffa's proof that if the axiom of counting held, then there was a permutation model in which no self-membered set could be finite. Friederike Körner and I both noticed that the assumption could be weakened to $\text{AxCount}_\leq$, and then further noticed that it could be weakened even to the existence of a function $f : \mathbb{N} \to \mathbb{N}$ such that $(\forall n)(n \leq f(Tn))$.

The result of the second of these two weakening steps can also be seen instead as the result of a *mild* mutagen experiment. The mild mutagen experiment first replaces selected occurrences of 'α' (not by '$T\alpha$' but) by '$f(T\alpha)$', and then prefixes '$(\exists f)$' at the start of the formula. So for example the mutagen experiment that took '$(\forall n \in \mathbb{N})(n \leq n)$' and gave us '$(\forall n \in \mathbb{N})(n \leq Tn)$' has a mild version that takes '$(\forall n \in \mathbb{N})(n \leq n)$' and gives us '$(\exists f : \mathbb{N} \to \mathbb{N})(\forall n \in \mathbb{N})(n \leq f(Tn))$'. If $\text{AxCount}_\leq$ fails, then T^{-1} is a very fast-growing function indeed, and on the face of it might grow faster than any function whose graph is a set. The weak version of $\text{AxCount}_\leq$ says that there are enough sets to prevent this happening, so that $T^{-1} \in O(f)$ for some function f whose graph is a set. The conjectures arising in this way are much weaker: '$(\forall n \in \mathbb{N})(n \leq Tn)$' appears to be quite strong, but the mild version '$(\exists f : \mathbb{N} \to \mathbb{N})(\forall n \in \mathbb{N})(n \leq f(Tn))$' is consistent relative to NF. A mutagen experiment turns $(\forall\alpha)(\alpha \leq \alpha)$ into $(\forall\alpha)(\alpha \leq T\alpha)$ which is refutable, but the mild version gives us $(\exists f : NO \to NO)(\forall\alpha)(\alpha \leq f(T\alpha))$ whose status is unclear.

Körner not only saw the use to which these functions $f : \mathbb{N} \to \mathbb{N}$ s.t. $(\forall n \in \mathbb{N})(n \leq f(Tn))$ could be put but also ([**5**]) showed how to obtain models of NF containing them. Accordingly functions $f : \mathbb{N} \to \mathbb{N}$ such that $(\forall n)(n \leq f(Tn))$ will here be called **Körner functions**, as will their generalisations to larger initial segments of the ordinals.

These two flavours of "Good-behaviour-of-T" conjecture will be mirrored elsewhere, as we will see.

Existence of large wellfounded sets. Let $\mathcal{P}_\kappa(X)$ be $\{Y \subseteq X : |Y| < \kappa\}$. κ does not have to be an aleph for this definition to make sense, but in all cases here it will be; indeed it will be a *cantorian* aleph. H_κ is the collection of sets hereditarily of size less than κ. We are not assuming foundation (this is NF after all), so we need to be clear that we mean here the *wellfounded* sets hereditarily of size less than κ. Thus

$$H_\kappa =: \{x : (\forall X)((\mathcal{P}_\kappa(X) \subseteq X) \to x \in X)\}$$

and we will be interested in the assertion "$\exists H_\kappa$" that H_κ exists, and in the assertion "$\Diamond\exists H_\kappa$" that there is a permutation model in which H_κ exists.

Now, since NF lacks unstratified separation we cannot be sure of the existence of the partition of H_κ into levels according to (set-theoretic) rank, *even if H_κ exists*. It is this partition that we need, or rather, we need what one might call the *segregation*: the cumulative family of unions of initial segments of the wellordered partition. We will notate this Π_κ.

The natural way to define this would be $\Pi_\kappa =: \{V_\alpha \cap H_\kappa : \alpha \in On\}$ but an obvious drawback is that although this represents our desired set as the image of On in a function that function is not obviously well-defined. We need a direct inductive definition. We say that Π_κ is the $\subseteq$-least collection X containing $\emptyset$, containing $\mathcal{P}_\kappa((x)$ whenever it contains x, and closed under unions, minus its last element (which is H_κ)

$$\Pi_\kappa := \{x : (\forall X)((\emptyset \in X \wedge \mathcal{P}_\kappa``X \subseteq X \wedge \bigcup``(\mathcal{P}(X)) \subseteq X) \to x \in X)\} \setminus \{H_\kappa\}$$

It turns out that it is the existence of Π_κ (rather than of H_κ) that has natural equivalences in ordinal-arithmetic-with-T. Let us abbreviate to '$\exists\Pi_\kappa$' the assertion that Π_κ exists.

The connection with arithmetic arises from $\mathrm{otp}(\Pi_\kappa)$, the length of Π_κ (in its natural order), or (which is the same if κ is cantorian) the supremum of the ranks of the sets hereditarily of power less than κ: we will be interested in the generalisation of $\mathrm{AxCount}_\leq$ to the ordinals below $\mathrm{otp}(\Pi_\kappa)$.

The fact that there is this obvious *prima facie* difference in NF between the strength of $\exists H_\kappa$ and $\exists\Pi_\kappa$ may not be as bizarre as it first appears. The existence of Π_κ is equivalent to the existence of the (set-theoretic) rank function restricted to H_κ and it is a well established (but not widely known) fact that there are facts about wellfounded sets that can be proved by induction on rank but not—apparently—by induction on $\in$. The most natural example of this is the phenomenon that from the assumption that the power set of any wellordered set can be wellordered one can prove by induction that every set is wellordered, and this induction is on rank not on $\in$.

No gratuitous failures of $\in$-foundation. We have already seen how Scott showed that there need be no Quine atoms. There is a sense in which this is an attractive result, in that although NF clearly proves the existence of illfounded sets—$V \in V$ is inescapable after all—there is no obvious reason why there should be *small* sets that violate foundation. Can one rein in failure of foundation even further than Scott did? We can get rid of self-membered singletons; can we perhaps get rid of self-membered finite sets? I raised this question in [**2**], and the answer turned out to be 'yes'. The path to this solution was shown us by Boffa, by means of a style of permutation which we will see below.

Now "no self-membered finite sets" is just a special case of "no $\in$-loops of length n involving only finite sets" and that in turn is just a special case of "$\in$ restricted to finite sets is wellfounded", and finiteness is merely one notion of smallness among others. So we have been led to the family of conjectures: "$\in$ restricted to small sets is wellfounded".

It turns out that there is a connection between "$\in$ restricted to small sets is wellfounded" and another suite of conjectures about NF, but to see it we have to turn back briefly from our headlong rush to the generality of "$\in$ restricted to small sets is wellfounded" and branch off at the point where we passed "no $\in$-loops of length n involving sets of size less than k" (where k and n are concrete natural numbers) for these conjectures are syntactically quite special. Let us say a formula in the language of set theory is $\exists^*$ if, once coerced into prenex normal form, all its quantifiers are existential, and $\forall^*\exists^*$ if, once coerced into prenex normal form, its quantifier prefix consists of universal quantifiers followed by existential. (That is, we do not give restricted quantifiers special treatment). Hinnion showed many years ago that NF proves every consistent $\exists^*$ formula. This result has been refined and improved in a variety of ways, and it is natural to wonder how many of these results apply also to $\forall^*\exists^*$ formulæ in NF. We cannot conjecture that NF proves every consistent $\forall^*\exists^*$ formula, since (as Scott showed in [**6**]) the $\forall^1\exists^1$ sentence '$(\forall x)(x \neq \{x\})$' is independent of NF. However there seems no obvious obstruction to a conjecture that

(1) If ϕ is a stratified $\forall^*\exists^*$ formula consistent with NF then NF $\vdash \phi$; and
(2) If ϕ is an unstratified $\forall^*\exists^*$ formula consistent with NF then NF $\vdash \Diamond\phi$.

There are some partial positive results and—so far—not even the faintest hint of counterexamples. The significance of this conjecture here is that many "no gratuitous failures of $\in$-foundation" assertions are unstratified $\forall^*\exists^*$ formulæ (for example: "no $\in$-loops of length n involving sets of size less than k" [where k and n are concrete natural numbers as before]) or $\forall^*\exists^*$ schemes, such as "No $\in$-loops involving only finite sets".

We have to be quite careful how to formulate conjectures embracing the idea that there should be no gratuitous failures of foundation, for $\{V\}$ is a small set that violates foundation. However it is not *hereditarily* small, so we might toy with conjectures like "every hereditarily small set is wellfounded", or perhaps "no finite set is self-membered". However, the strongest natural conjectures suggested by this line of thought are things like "$\in$ restricted to small sets is wellfounded" (for some suitable notion of smallness): perhaps every collection with no $\in$-minimal element contains a large set.

A little further thought reveals that sentences like "$\in$ restricted to small sets is wellfounded" admit the same kind of strengthening that we saw in the preceding section, namely the strengthening that took us from "$\exists H_\kappa$" to "$\exists \Pi_\kappa$". This strengthens "$\in$ restricted to small sets is wellfounded" to "the graph exists of the rank function of $\in$ restricted to small sets" or—equivalently—to "Δ_κ exists". Here Δ_κ is inductively defined as the $\subseteq$-least collection X containing $\emptyset$ and—if we write $b_\kappa(x)$ for $\{y : |y| < \kappa \wedge (\forall z \in y)(|z| < \kappa \rightarrow z \in x\}$—containing $b_\kappa(x)$ whenever it contains x, and closed under unions. Δ_κ is naturally wellordered by $\subseteq$: its bottom element is $\emptyset$; then comes the set of all sets of size $< \kappa$ that have no members of size $< \kappa$; then comes the set of those sets of size $< \kappa$ all of whose members of size $< \kappa$ have already appeared; ....

We let $\exists\Delta_\kappa$ be the assertion that Δ_κ exists.

1. The results

Consider the seven assertions drawn from these three areas of concern.

(1) $\exists H_\kappa$: H_κ exists;
(2) $\exists \Pi_\kappa$: Π_κ exists;
(3) $\langle \{x : |x| < \kappa\}, \in\rangle$ is wellfounded;

(4) $\exists\Delta_\kappa$: Δ_κ exists;
(5) $(\forall\alpha < \kappa^*)(\alpha \leq T\alpha)$;
(6) Sizes of wellfounded sets are not bounded below κ;
(7) There is a κ-Körner function: $f : \{\alpha : \alpha < \kappa\} \to \{\alpha : \alpha < \kappa\}$ s.t $(\forall\alpha < \kappa)(\alpha \leq f(T\alpha))$.

The precise meaning of κ^* (in 5) will become clear as we go on. It's either κ or κ^+. In what follows we will assume without further comment that κ is a cantorian aleph.

Throughout the above one can replace '$|x| < \kappa$' with '$|x| \not\geq^* \kappa$' and—although on the whole it doesn't make much difference to the implications between the formulæ above—it does make the inference of (the modified version of) 6 from (the modified version of) 1 slightly easier, as follows. H_κ is wellfounded, and so is not self-membered. But it is a set of wellfounded sets hereditarily of size $< \kappa$ so the only way it can avoid being hereditarily of size $< \kappa$ itself is to be of size $\not< \kappa$. But then one does not need much extra information to infer that it has subsets of all sizes below κ. This inference becomes easier if we consider instead H^*_κ: the set of all wellfounded sets hereditarily too small to map onto a set of size κ, for then we know that there is a wellfounded set (to wit: H^*_κ) that has a partition of length κ.

On the whole we will not pursue the extra generality to be achieved by extending proofs to the $|x| \not\geq^* \kappa$ case.

The following implications are all obvious.

$2 \to 1$, $4 \to 3$, $4 \wedge 1 \to 2$ and $5 \to 7$

These implications are not obvious, and will be proved below:

$\Diamond 2 \to 5$; $\Diamond 2 \to \Diamond$ 4; $\Diamond 4 \to 5$; $5 \to \Diamond 2$; $\Diamond 1 \to \Diamond 3$. $6 \to \Diamond 3$ holds if κ is strongly inaccessible; $7 \to \Diamond 3$ holds as long as κ is regular.

Further strengthenings of these implications arise from the fact that $\Diamond\Diamond A \to \Diamond A$, so whenever we have proved $A \to \Diamond B$ then we have also proved $\Diamond A \to \Diamond B$.

The burden of the preceding discussion is that one would expect 1- 7 to divide into two bundles of equivalent propositions, namely $\{\Diamond 1, \Diamond 3, 7\}$ and $\{\Diamond 2, \Diamond 4, 5\}$, corresponding to the mild and to the basic mutagen experiment respectively, and where each bundle contains one representative from each of the three classes of proposition in the list on page 2. However a number of implications remain unproved for the moment.

Some notation

Our ordered pairs will be Quine ordered pairs, and we will write them with angle brackets: $\langle x, y\rangle$. (The chief feature of Quine ordered pairs that concerns us is that $|\langle x, y\rangle| = |x|+|y|$; other details can be found in [2]). $\texttt{fst}(x)$ and $\texttt{snd}(x)$ are the first and second components of the ordered pair x. Round brackets—(x, y)—denote the transposition swapping x with y, so

$$\prod_{\Phi(x,y)} (x, y)$$

is the product of all transpositions that swap x with y as long as $\Phi(x, y)$, and fixing everything else. (This notation can be safely used only when all these transpositions are disjoint, and this disjointness will need to be argued for in each case).

1.1. Boffa permutations. The chief tool we use here to obtain relative consistency results is Rieger-Bernays permutation models, and specifically a family of permutations whose prototypical member was discovered by Boffa [**1**] and used by him to give a partial solution to my question "Is it consistent relative to NF that no finite set can be a member of itself?" It is the mark of good mathematicians that they have good ideas, and it is the mark of good ideas that they outrun the purposes for which they were dreamed up. We will now never know how much of what we now know we can prove with Boffa permutations was foreseen by Boffa, for he is no longer with us to explain.

A Boffa permutation has four ingredients:

(1) A notion of smallness. This will be a property whose extension is closed under subsets and surjective images, so that in particular π is a permutation of the universe, and x is small, so is $\pi``x$. This last will ensure that, for any permutation π, V^π believes x to be small if and only $\pi(x)$ is small. We also want complements of small things to not be small, and finite unions of small things to be small. Finally, since we are using Quine ordered pairs, we find that the foregoing implies that an ordered pair is small iff both its components are small (since $|\langle x, y\rangle| = |x| + |y|$ as remarked earlier).

(2) A structure $\langle S, E\rangle$ where E is (i) stratified-but-inhomogeneous in the sense that $\{\langle\{x\}, y\rangle : x\ E\ y\}$ is a set, and (ii) wellfounded in the sense that every subset of S has an E-minimal element.

(3) A function f from $\mathcal{P}_{\text{small}}(S)$ (the small subsets of S) to S such that whenever $x \in X$ and X is a small subset of S then $x\ E\ f(X)$. We want this f to be homogeneous so that its graph is a set. f is not required to be injective.

Finally, we have not so far considered whether the elements of S themselves—considered as sets—are small or not. It doesn't matter one way or the other (as long as they are all small or all not small) but we have to know which it is. If elements of S are not small then we set π to be the permutation

$$\prod_{\text{small}(x)} (x, \langle x, f((\texttt{snd}``x) \cap S)\rangle)$$

that swaps every small x with $\langle x, f((\texttt{snd}``x) \cap S)\rangle$. On the other hand if elements of S are small we set π to be the permutation

$$\prod_{\text{small}(x)} (x, \langle -x, f((\texttt{snd}``x) \cap S)\rangle)$$

The idea is that π shall always swap small things with things that are not small. It is elementary to check that all these transpositions are disjoint, and that the product is well-defined.

The following lemma is the means of application of Boffa permutations.

LEMMA 1.1. *If we write '$SMALL^\pi$' for the collection of those sets which V^π believes to be small, then* $\texttt{snd}$ *is a homomorphism from $\langle SMALL^\pi, \in^\pi\rangle$ to $\langle S, E\rangle$.*

Proof: There are two cases to consider, depending on whether or not elements of S are small. The argument is essentially the same in both cases, so we will treat only one. Let's assume that, as in the original case, elements of S are not small.

Suppose that x and y are both believed by V^π to be small, and that $x \in_\pi y$. We will prove $\texttt{snd}(x)\ E\ \texttt{snd}(y)$.

Since x and y are both believed by V^π to be small, $\pi(x)$ and $\pi(y)$ are both small. We have cooked up π so that $\langle x, f((\texttt{snd}``x) \cap S)\rangle$ cannot be small. (It's an ordered pair one of whose components is in $S$ and is accordingly not small.) Therefore, since $\pi(x)$ is small, $x$ must be not-small and must be $\langle \pi(x), f((\texttt{snd}``\pi(x)) \cap S)\rangle$. Similarly y must be $\langle \pi(y), f((\texttt{snd}``\pi(y)) \cap S)\rangle$.

Now $x \in \pi(y)$ so $\texttt{snd}(x) \in \texttt{snd}``(\pi(y))$. Also $\texttt{snd}(x) = f(\texttt{snd}``(\pi(x)) \cap S)$, which is in S (being a value of f), so $\texttt{snd}(x)$ is in $\texttt{snd}``(\pi(y)) \cap S$. Now $\texttt{snd}``(\pi(y)) \cap S$ is a small subset of S, and so an argument to f. But now (since $x\ E\ f(X)$ whenever $x \in X$ and X is a small subset of S) we can infer $\texttt{snd}(x)\ E\ f(\texttt{snd}``(\pi(y)) \cap S)$. Finally $f(\texttt{snd}``(\pi(y)) \cap S) = \texttt{snd}(y)$ so $\texttt{snd}(x)\ E\ \texttt{snd}(y)$ as desired.

Thus we have shown that: whenever V^π thinks that $x \in y$ and that both x and y are small, then $\texttt{snd}(x)\ E\ \texttt{snd}(y)$, and we also know that both of these things are in S.

■

The effect of this is to ensure that V^π believes that $\in$ restricted to small sets is wellfounded, (every set of small sets has an $\in$-minimal element) since there is a homomorphism from the small sets of V^π onto $\langle S, E\rangle$

(The alert reader may have spotted that we could have achieved the same effect by swapping our ordered pairs round and using `fst` instead of `snd`. The extra flexibility will not be exploited here, but one hopes it will be one day.)

In the original case discovered by Boffa the notion of smallness was finitude; S was $\mathbb{N}$; E was $\{\langle\{n\}, m\rangle : Tn < m\}$; and $f(X)$, for X a finite set of natural numbers, was $T(\inf(\mathbb{N} \setminus X))$. Boffa assumed the axiom of counting, with which he could prove that E is wellfounded. ($\text{AxCount}_\leq$ is in fact sufficient). I do not believe he ever published this result. Using $T(\sup(X))$ instead of $T(\inf(\mathbb{N} \setminus X))$ is the crucial change that gives results about wellfoundedness instead of mere self-membership.

1 implies $\Diamond$3 and 2 implies $\Diamond$4

Let us start with a simple application of the machinery we have just seen. In fact the same permutation will do for both implications.

For $1 \to \Diamond 3$ take $\langle S, E\rangle$ to be $\langle H_\kappa, \in\rangle$, take smallness to be $|x| < \kappa$, and f to be the identity. Then things in S are small, and we need the flavour of Boffa permutation where we complement the first component of the ordered pair. Then we have a straightforward instance of lemma 1.1.

It is slightly more work to verify that the resulting permutation model satisfies 4 if we start with one that satisfies 2. Let σ be the permutation we have just characterised. It will be sufficient to locate the set which in V^σ is Δ_κ. This set must be

$$\{\{y : \texttt{snd}``y \cap (\bigcup \Pi_\kappa) \in x\} : x \in \Pi_\kappa\}$$

$\Diamond$2 is equivalent to 5

First we need some definitions, in order to redeem the pledge of κ^*.

We will be interested in isomorphism types of structures called BfExts. What is a BfExt? The word is an abbreviation of 'Bien fondée extensionnelle'. If x is a wellfounded set, then $\langle TC(\{x\}), \in, x\rangle$ is the prototypical example of a BfExt. (At least if $\in |TC(\{x\})$ exists!) That is to say a BfExt is a binary structure $\langle X, R, x^*\rangle$ with a binary relation R and a distinguished element x^* satisfying the conditions that

(1) R is wellfounded and extensional;
(2) x^* is the unique $x \in X$ such that $t(R)``\{x\} = X \setminus \{x\}$.

$t(R)$ is the transitive closure of R. x^* is the "top" element: x is the "top" element of $\langle TC(\{x\}), \in, x\rangle$.

We also need a binary relation between isomorphism classes of BfExts. This is the one that is motivated by the relation that holds—in ordinary set theory—between the two BfExts $\langle TC(\{x\}), \in, x\rangle$ and $\langle TC(\{y\}), \in, y\rangle$ when $x \in y$. We say $\langle X, R, x^*\rangle \boldsymbol{\in} \langle Y, S, y^*\rangle$ when there is an element $y^\dagger$ of Y such that $\langle y^\dagger, y^*\rangle \in S$ and the BfExt obtained from $\langle Y, S, y^*\rangle$ by cutting Y down to $\{y \in Y : \langle y, y^\dagger\rangle \in t(S)\}$ ($t(S)$ is the transitive closure of S), restricting S to this set, and taking $y^\dagger$ as the new top point, is isomorphic to $\langle X, R, x^*\rangle$. We will overload '$\boldsymbol{\in}$' to mean both the relation between BfExts and the induced relation between the isomorphism types of BfExts.

It is an important elementary fact that $\boldsymbol{\in}$ is wellfounded and so has a rank function, and we write 'ρ' for that rank function.

BfExts are interesting in general, but here we will be specifically interested in the set $\mathfrak{K}$ of relational types of what one could call "κ-like BFexts', where every element has fewer than κ "immediate predecessors". When $x \in H_\kappa$, $\langle TC(\{x\}), \in, x\rangle$ is the prototypical example of a κ-like BfExt. To be precise: $\langle X, R, x^*\rangle$ is κ-like when, for all $x \in X$, $|R^{-1}``\{x\}| < \kappa$.

THEOREM 1.2. *Let κ be a cantorian aleph, $\mathfrak{K}$ the set of relational types of κ-like BfExts, and κ^* the rank of $\langle \mathfrak{K}, \boldsymbol{\in}\rangle$. Then the following are equivalent*

$$(\forall \alpha < \kappa^*)(\alpha \leq T\alpha); \qquad 5$$
$$\Diamond \exists \Pi_\kappa. \qquad \diamond 2$$

Proof:

5 **implies $\Diamond$2.** In ([**4**]) Jech showed—without any use of AC—that every element of $H_{\aleph_1}$ has rank $< \omega_2$. We here spice up the elegant argument he used there, and use it to establish that κ^* is defined and is κ^+ at most.

LEMMA 1.3. *Everything in $\mathfrak{K}$ is of rank $< \kappa^+$.*

Proof:

We define a function $F : NO^{<\omega} \to \mathfrak{K} \to NO$ as follows, using `ordlist` as a variable for lists of ordinals and '::' for the operation of `cons`ing an ordinal onto the front of a list of ordinals. [] is the empty list.

$$F \quad [\ ] \quad x =: \rho(x);$$
$$F \quad (\alpha::\texttt{ordlist})\ x =: \alpha^{th} \text{ member of } \{(F\ \texttt{ordlist}\ y) : y \boldsymbol{\in} x\}$$
$$\text{or } 0, \text{ if this is undefined.}$$

Let K be the set of ordinals below κ and identify K^n with the set of lists of length n whose elements are ordinals below κ. We will need the following observation:

LEMMA 1.4. *For $n \in \mathbb{N}^+$ and $x \in \mathfrak{K}$, every element of $\{\rho(y) : y \in^n x\}$ is a value for argument x of one of the functions in $F``K^n$.*

Proof: When $n = 1$ we observe that, since $x \in \mathfrak{K}$ there are α ordinals that are ranks of members$^{\in}$ of x for some $\alpha < \kappa$ so every one of them is $F\ [\alpha]\ x$ for some suitable α.

In the general case suppose $x_n \in x_{n-1} \in \ldots x_0 = x$. We want to find a list $[\alpha_n, \alpha_{n-1}, \ldots]$ of ordinals below κ such that $\rho(x_n) = F\ [\alpha_n, \alpha_{n-1}, \ldots]\ x$.

α_n must clearly be that α such that $\rho(x_n)$ is the αth member of $\{\rho(y) : y \in x_{n-1}\}$, so $F\ [\alpha_n]\ x_{n-1} = \rho(x_n)$, and x_{n-1} is the argument.

By the same token α_{n-1} must be that α such that $(F\ [\alpha_n]\ x_{n-1}$ is the αth member of $\{(F\ [\alpha_n]\ y) : y \in x_{n-2}\}$, so $F\ [\alpha_n; \alpha_{n-1}]\ x_{n-2} = \rho(x_n)$, and x_{n-2} is the argument.

Again α_{n-2} must be that α such that $F\ [\alpha_n; \alpha_{n-1}]\ x_{n-2}$ is the αth member of $(F\ [\alpha_n; \alpha_{n-1}]\ y) : y \in x_{n-3}\}$, so $F\ [\alpha_n; \alpha_{n-1}; \alpha_{n-2}]\ x_{n-3} = \rho(x_n)$, and x_{n-3} is the argument.

The subscript on the argument gets smaller at each stage...

This completes the proof of lemma 1.4. ■

This algorithm for calculating F assigns a list of ordinals below κ to every finite sequence $x_n \in x_{n-1} \in \ldots x$. So for every $x' \in^n x$ there is at least one (and possibly many) such lists corresponding to it. Clearly for fixed x the function $\lambda l.(F\ l\ x)$ is total (there is a failure-trapping clause after all), and so it is a surjection from $K^{<\omega}$ onto $\{\rho(y) : ((\exists n)(y \in^n x)\}$. So $\kappa = |K^{<\omega}| \geq^* |\{\rho(y) : (\exists n)(y \in^n x)\}| = |\rho(x)|$. That is to say κ is at least as big as the initial ordinal corresponding to $\rho(x)$. In other words, for $x \in \mathfrak{K}$, $\rho(x) < \kappa^+$, as we claimed.

This completes the proof of lemma 1.3. ■

There is a T operation on BfExts, and T of a κ-like BfExt is another κ-like BfExt (because $\kappa = T\kappa$). We establish by induction on E that T commutes with ρ.

We now know that κ^* is defined, so let us assume $(\forall \alpha < \kappa^*)(\alpha \leq T\alpha)$. This is enough to show that the relation $T\alpha \in \beta$ on relational types of κ-like BFexts is wellfounded, as follows. Suppose X were a subset of $\mathfrak{K}$ s.t. for all $x \in X$ there is $y \in X$ with $Ty \in x$. Now let α be the least ordinal in $\rho``X$. $\alpha = \rho(x)$ for some $x \in X$ and there is $y \in X$ with $Ty \in x$. So $\rho(y) \geq \alpha > \rho(Ty) = T\rho(y)$. So $\alpha > T\alpha$ contradicting $(\forall \alpha < \kappa^*)(\alpha \leq T\alpha)$.

Now we can exhibit the permutation which will make $\Diamond\exists\Pi_\kappa$ true. Let σ be

$$\prod_{\alpha \in \mathfrak{K}} (T\alpha, \{\beta : \beta \in \alpha\}).$$

The product is well-defined since the transpositions are all disjoint: nothing is both a κ-like BfExt and a set of κ-like BfExts. In V^σ $\mathfrak{K}$ has become a set of sets hereditarily of size less than κ and, because $(\forall \alpha < \kappa^*)(\alpha \leq T\alpha)$, it is precisely the sets of *wellfounded* sets hereditarily of size less than κ, which is to say H_κ.

Finding Π_κ in V^σ is slightly harder. The critical fact here is that the prewellordering of sets according to set-theoretic rank (x related to y iff $\rho(x) \leq \rho(y)$ where ρ here is—for once, confusingly—set-theoretic rank as usual) is the least fixed point for the '+' operation on quasiorders, defined so that, for a quasiorder $\leq$, $X \leq^+ Y$ iff $(\forall x \in X)(\exists y \in Y)(x \leq y)$.

The set which will be Π_κ in V^σ is the cumulative version of the partition of $\mathfrak{K}$ according to the rank function for $T \circ \in$ (the relation $\{\langle x, y\rangle : Tx \in y\}$). We certainly have the partition of $\mathfrak{K}$ according to the rank function for $\in$, and it turns out that this is the same partition.

We establish this by proving that the prewellordering-according-to-ρ (which is a set and which we write $\leq_\rho$)

(i) is a superset of $T \circ \in$;

(ii) is a pre-fixed point for the $\in$-version of the '+' operation (where $x\ R^+\ y$ iff $(\forall x')(Tz \in x \to (\exists y')(Ty \in y \wedge x'\ R\ y')))$ and

(iii) is a subset of all other pre-fixed points.

We prove these as follows.

(i) $Tx \in y$ implies $\rho(Tx) < \rho(y)$. T commutes with $\in$ so this implies $T\rho(x) < \rho(y)$. But we are assuming $\rho(x) \leq T\rho(x)$ so we infer $\rho(x) < \rho(y)$ as desired.

(ii) We must show that if $(\forall x')(Tz \in x \to (\exists y')(Ty' \in y \wedge \rho(x') \leq \rho(y'))$ then $\rho(x) \leq \rho(y)$. Now T permutes $\mathfrak{K}$ so the assumption is just $(\forall x')(z \in x \to (\exists y')(y' \in y \wedge \rho(x') \leq \rho(y'))$ which is the recursion that defined ρ in the first place.

(iii) Let R be a prefixed point for $+$. We will prove by induction on $\in$ that it is a superset of $\leq_\rho$. Let x be $\in$-minimal such that there is y with $x \leq_\rho y$ but $\neg(x\ R\ y)$, and let y be $\in$-minimal for x. Then there is $x' \in x$ such that for no $y' \in y$ do we have $x'\ R\ y'$. But for any $x' \in x$ whatever there is $y' \in y$ such that $x' \leq_\rho y'$ whence $x'\ R\ y'$ by minimality of x so $x\ R^+\ y$ and $x\ R\ y$ because R is a prefixed point.

■

Now for the other direction of theorem 1.2.

$\Diamond 2$ **implies** 5. (This result follows from the facts that $\Diamond 2 \to \Diamond 4$ and $\Diamond 4 \to 5$, but we retain the proof here since it might be felt to be of independent interest, and helps to prepare us for the proof of $\Diamond 4$ implies 5, which is closely analogous to it.)

Assume 2 (so that H_κ and Π_κ both exist) and let $H : \{\alpha \in NO : \alpha < \kappa^*\} \to \Pi_\kappa$ enumerate Π_κ in its natural order, under $\subseteq$. κ^* is the length of Π_κ. The assertion

$$(\forall \alpha)(\mathcal{P}_\kappa(H(\alpha)) = H(T\alpha + 1)) \tag{1.1}$$

is weakly stratified and can be proved by an induction on α. Now suppose *per impossibile* there were $\alpha < \kappa^*$ with $\alpha > T\alpha$. Then we would have

$$\mathcal{P}_\kappa(H(\alpha)) = H(T\alpha + 1) \tag{1.2}$$

Now $\alpha > T\alpha$ implies $\alpha > T\alpha + 1$ since $\alpha = T\alpha + 1$ is impossible. This gives $H(T\alpha + 1) \subseteq H(\alpha)$ and consequently $\mathcal{P}_\kappa(H(\alpha)) \subseteq H(\alpha)$. So the process has closed at a stage $< \kappa^*$, contradicting the definition of κ^* as the closure ordinal of this construction.

So we have assumed $\exists \Pi_\kappa$ and proved $(\forall \alpha < \kappa^*)(\alpha \leq T\alpha)$. But the conclusion is invariant so it follows even from $\Diamond \exists \Pi_\kappa$. ■

$\Diamond 4$ implies 5

We don't mean *exactly* that $\Diamond 4$ implies 5: the bound that appears in 5 is an ordinal defined rather in the way that κ^* (from 2) was defined. The proof is analogous to the proof of $\Diamond 2$ implies 5 that we have just seen.

Assume 4, and let $A : \{\alpha \in NO : \alpha < \lambda\} \to \Delta_\kappa$ enumerate Δ_κ in its natural order, under $\subseteq$. (λ is the closure ordinal, that stage at which b_κ ceases to give any new sets of size less than κ). We prove that

$$(\forall \alpha)(A(T\alpha + 1) = b_\kappa(A_\alpha)) \tag{1.3}$$

by a careful induction as in the analogous case (formula 1.1 above) with $\Diamond \exists \Pi_\kappa$.

Now suppose there is $\alpha < \lambda$ with $\alpha > T\alpha$. Then we would have

$$b_\kappa(A(\alpha)) = A(T\alpha + 1) \tag{1.4}$$

by analogy with formula 1.1. Now $\alpha > T\alpha$ implies $\alpha > T\alpha + 1$ since $\alpha = T\alpha + 1$ is impossible as before. This gives $A(T\alpha + 1) \subseteq A(\alpha)$ so $b_\kappa(A(\alpha)) \subseteq A(\alpha)$ and the process has closed at a stage $< \lambda$ contradicting the definition of λ as the closure ordinal of this construction.

So we have assumed 4 and proved $(\forall \alpha < \lambda)(\alpha \leq T\alpha)$. But the conclusion is invariant so it follows even from $\Diamond 4$.

If κ is regular then 7 implies $\Diamond$ 3

We use a Boffa permutation. Smallness is "smaller than κ". There is a Körner function g; we take S to be the ordinals below κ; E to be $\{\langle\{\alpha\}, \beta\rangle : g(T\alpha) < \beta\}$; finally $f(X)$, for X a set of ordinals below κ, is $T(\sup(X))$. The assumption that κ is regular is needed to ensure that f is well-defined.

Elements of S are not themselves small, so we use the first flavour of π, the one that does not complement the first component of the ordered pair. ■

If κ is strongly inaccessible then 6 implies $\Diamond$ 3

Let κ be a strongly inaccessible cantorian aleph, and suppose that there are wellfounded sets of arbitrarily large size below κ. We know anyway that every self-membered power set contains all wellfounded sets, and so—in this setting—cannot be of size $< \kappa$. Again, we know anyway that, if B is a collection of power sets with no $\in$-minimal member, then $\bigcap B$ is a self-membered power set. So nothing in B can have size less than κ. So $\in$ restricted to $\{\mathcal{P}(x) : |x| < \kappa\}$ is wellfounded, and we then use a Boffa permutation where S is $\{\mathcal{P}(x) : |x| < \kappa\}$; E is the restriction to S of $\{\langle\{x\}, y\rangle : x \in y\}$; smallness is $|x| < \kappa$; finally $f(X)$, for X a small set of small power sets, is $\mathcal{P}(\bigcup X)$. Things in S are small so we use the *second* flavour of Boffa permutation, where we complement the first component of the ordered pair. ■

2. Open Problems

There are various loose ends that later workers may tie up. The world would be a pleasantly tidy place indeed if 1 - 7 were to divide (as indicated earlier) into two bundles of equivalent propositions, namely $\{\Diamond 1, \Diamond 3, 7\}$ and $\{\Diamond 2, \Diamond 4, 5\}$. We have shown that the three formulæ in the second bundle are equivalent, but several conjectured equivalences within the first bundle remain to be proved. In particular I cannot at this stage see any way of constructing Körner functions from large wellfounded sets. However, I am endebted to the referee for the observation that the methods of [**5**] will produce Körner functions on $\{\alpha \in NO : \alpha < \kappa\}$ where κ is a definable cantorian ordinal.

The extreme bottom and the extreme top of this scale (1-7) of propositions both merit special attention.

At the bottom, where $\kappa = \aleph_0$, we know that 7 is consistent relative to NF; for all we know $\Diamond 6$ and $\Diamond 3$ could be theorems of NF. This has always seemed unlikely, and now seems even more unlikely in the light of a recent result of Holmes (to appear) to the effect that no consistent invariant extension of NF proves the existence of infinite wellfounded transitive sets.

Although it is suspected that we can prove Con(NF) in the arithmetic of NF + $\text{AxCount}_{\leq}$ (2 becomes $\text{AxCount}_{\leq}$ when $\kappa = \omega$) this has never been proved, and $\text{AxCount}_{\leq}$ could be consistent relative to naked NF (though this seems highly unlikely).

It is still open whether or not $\text{AxCount}_{\leq}$ implies the analogue $5(\aleph_1)$ for countable ordinals.

At the top end (where κ is an aleph but no longer cantorian) one can ask: might the collection of hereditarily wellorderable sets be a set? What about the consistency strength of a Körner function over the whole of the ordinals? What about Körner functions that commute with T? There are presumably analogues of Körner functions for BfExts: do they give rise to any new mathematics?

Finally there are conjectures along the lines of 1-7 that arise if we consider notions of smallness not parameterised by a cardinal. Strongly cantorian is one such property, but there are others. Let a **small** set be one that does not map onto V and a **low** set one that is the same size as a wellfounded set.

The analogues of 6 are: Is every wellfounded set small? (Is the set of small sets itself small?) Is every wellfounded set strongly cantorian?

The analogues of 3 ask whether or not we can arrange for $\in$ restricted to small/low/strongly cantorian sets to be wellfounded. The standard methods of constructing models for Church's set theory (see Forster [**3**]) easily give models of fragments of NF in which $\in$ restricted to low sets is wellfounded, but the standard methods do not construct models of NF and we do not know at present how to obtain permutation models of NF in which $\in$ restricted to low sets is wellfounded. Boffa-Pétry [**1**] used a Boffa permutation to obtain a model in which there are no strongly cantorian self-membered sets. Their proof uses the axiom of counting, and although the result is unlikely to be best possible, I cannot at the moment see how to drop this assumption nor how to strengthen the conclusion to get a permutation making $\in$ restricted to strongly cantorian sets wellfounded.

Finally one can wonder whether there is a table like 1-7, but using $\Box$ instead of $\Diamond$. For example, we can prove the following

Remark 2.1. $\Box(\forall x)(|x| \in \mathbb{N} \to \mathcal{P}(x) \not\subseteq x)$ is equivalent to $\text{AxCount}_{\leq}$

Proof:

$\neg$R $\to \neg$ L: If $\text{AxCount}_{\leq}$ fails, there is $n > 2^{Tn} \in \mathbb{N}$. Since whenever $x \notin x$, $\{y : x \in y\}$ is a set of size $|V|$ disjoint from its power set, we can find, for any cardinal n, a set of size n disjoint from its power set. In particular if n is the finite cardinal promised above (so that $2^{Tn} < n$) then we have a set x of size n disjoint from its power set and an injection p from $\mathcal{P}(x)$ into x. This can be extended to a permutation π of V, and in V^π every wellfounded set has size $< n$. This proves $\Diamond\neg(\forall x)(|x| \in \mathbb{N} \to \mathcal{P}(x) \not\subseteq x)$.

R $\to$ L: If π is a permutation such that V^π thinks that some set x is finite and a superset of its power set, then V contains a map (namely a suitable restriction of π) from some finite power set $\mathcal{P}(x)$ into x and therefore a natural number $n = |x|$ such that $2^{Tn} < n$, which contradicts AxCount$_\leq$.

Moral

People who think that Set theory is ZFC will probably feel that the moral pointed by the equivalences announced in this paper is hardly news, and indeed was known already to Horace: *Naturem expelles furca, tamen usque recurret.* In seeking a resolution of the paradoxes, it matter not to what lengths you go to avoid the cumulative hierarchy, driving it away with a fork, it will come back out of the bush and bite your ankles.

But I think the true moral is more subtle and more interesting: what is coming back with a vengeance is not the cumulative hierarchy, but the mathematics itself. Your choice of formalisation will determine whether the solution to the paradoxes comes out as funny-ordinals-with-a-T-function, or as the cumulative hierarchy, or as the theory of types, or God-knows-what; but nothing can prevent these various formalisations from being mutually interpretable. In a deep sense there is only one solution.

Finally I would like to thank Randall Holmes and the anonymous referee for helpful comments.

References

[1] Boffa, Maurice. and Pétry, André. On self-membered sets in Quine's set theory NF. Logique et Analyse 141-142, (1993) pp. 59-60.
[2] Forster, Thomas. Set theory with universal set: exploring an untyped universe. second edition.
[3] Forster, Thomas. Church-Oswald models for Set Theory. in: Logic, Meaning and Computation: essays in memory of Alonzo Church, Synthese library **305**. Kluwer, Dordrecht, Boston and London 2001 (revised edition, *sans* missprints, available on `www.dpmms.cam.ac.uk/~tf/churchlatest.ps`).
[4] Jech, Thomas. On Hereditarily countable sets. Journal of Symbolic Logic **47** (1982) pp. 43-47.
[5] Körner, Friederike. Cofinal indiscernibles and some applications to New Foundations. Mathematical Logic Quarterly **40** (1994), pp. 347-356.
[6] Scott, Dana. Quine's individuals. Logic, Methodology and Philosophy of science, ed. E. Nagel, Stanford University Press (1962), pp. 111-115.

Department of Pure Mathematics and Mathematical Statistics
Centre for Mathematical Sciences
Wilberforce road
Cambridge
CB3 0WB
United Kingdom