INTERNAL AUTOMORPHISMS AND ANTIMORPHISMS OF MODELS OF NF

NATHAN BOWLER AND THOMAS FORSTER

Abstract. It is shown that every model of NF admits a permutation model containing an internal automorphism.

The dual $\phi^{\circlearrowleft}$ of a formula ϕ is the formula obtained from ϕ by replacing all occurrences of '$\in$' in ϕ by '$\notin$'. The axiom scheme $\phi \longleftrightarrow \phi^{\circlearrowleft}$ is the Duality Scheme. It has been known for some time that $\phi \longleftrightarrow \phi^{\circlearrowleft}$ is a theorem of NF whenever ϕ is a closed stratifiable formula (the $\circlearrowleft$ operation does not affect stratification). Permutation models can be found in which $\phi \longleftrightarrow \phi^{\circlearrowleft}$ fails for some unstratifiable ϕ, but it remains an open question whether or not there are models in which $\phi \longleftrightarrow \phi^{\circlearrowleft}$ holds for all ϕ. (The place to look for the details is the chapter on permutation models in [1], which also contains all the background that a reader might need for what follows below.) The natural conjecture is that there should be such models. An *antimorphism* is a permutation τ of V satisfying $(\forall x, y)(x \in y \longleftrightarrow \tau(x) \notin \tau(y))$. Clearly if there is an antimorphism then duality follows (tho' one does not expect a converse, since the existence of antimorphisms of order two contradicts AC_2).

We do not prove the conjecture here, but we do prove a special case.

We say a formula ϕ is *stratifiable-mod-2* if its variables can be assigned to two types `yin` and `yang` in such a way that:

(i) all occurrences of any one variable receive the same type, and

(ii) in subformulæ like '$x = y$' the two variables receive the same type, and

(iii) in subformulæ like '$x \in y$' the two variables receive different types.

In Corollary 3 we establish that every model of NF has a permutation model satisfying the scheme $\phi \longleftrightarrow \phi^{\circlearrowleft}$ for ϕ that are stratifiable-mod-2. As a side-effect of our analysis we obtain a proof that every model of NF (and not just "every model of NF + AC_2" which was hitherto the best known) has a permutation model containing an $\in$-automorphism that is both nontrivial and internal (a set of the model). This is Corollary 2.

> We record for clarity that all proofs are conducted in NF (not ZF) and indeed NF *simpliciter*, with no add-ons. Readers more at home

Received July 2, 2023.
2020 *Mathematics Subject Classification.* Primary 03Exx.
Key words and phrases. antimorphism, duality.

0022-4812/00/0000-0000
DOI:10.1017/jsl.2024.47

with ZF might need to remind themselves that in this context 0 and $\emptyset$ are not the same! $\emptyset$ is always the empty set, but 0 is the natural number zero.

We will need to consider the sequence of permutations: $\mathbf{1}$, c, $jc \cdot c$, $j^2c \cdot jc \cdot c$ ..., where c is the complementation permutation, and the operator j is defined so that $j(\pi)$ is the function $x \mapsto \pi``x$. We write these permutations 'c_i', thus: $c_1 := c$; $c_{i+1} := j(c_i) \cdot c$.

Definition 1. For permutations σ and τ of sets X and Y, an embedding of permutations from σ to τ is an injective function $\pi\colon X \to Y$ such that $\pi \cdot \sigma = \tau \cdot \pi$.

Although Definition 1 is quite general we will need it in this paper only for involutions, and we will speak of involution-embeddings or embeddings of involutions. (A permutation π is an involution iff $\pi^2 = \mathbf{1}$.)

We will need the following analogue of Cantor–Bernstein for embeddings-of-involutions.

Lemma 1. *Let σ and τ be involutions of X and Y such that there are embeddings π of σ into τ and ρ of τ into σ.*
Then σ and τ are conjugate.

Proof. Most proofs of the Cantor–Bernstein theorem extend to proofs of this fact. For the sake of brevity, we will use a proof based on the Knaster–Tarski theorem that any order-preserving function on a complete lattice has a fixed point. Applying this to the lattice of sets which are closed under the action of σ and the order-preserving function $S \mapsto X \setminus \rho``(Y \setminus \pi``S))$ we obtain a fixed point P. Then the map defined by π on P and ρ^{-1} on $X \setminus P$ is an isomorphism from σ to τ. ⊣

We observe without proof that if π is an embedding of permutations from σ to τ then $j(\pi)$ is an embedding of permutations from $j(\sigma)$ to $j(\tau)$. That is to say, conjugacy is a congruence relation for j, so we can think of j as acting on the congruence classes. (We will need this in the proof of the second part of Lemma 2.)

Definition 2. An involution is universal if every involution embeds into it.

Lemma 2. *For all i, $j^i(c)$ is universal.*

Proof. First we prove that $j(c)$ is universal.

We are going to need a bijection from V to $\{x : \emptyset \notin x\}$. First we define $f : V \to V$ by

$$x \mapsto \begin{cases} x+1, & \text{if } x \in \mathbb{N} \setminus \{0\}, \\ 1, & \text{if } x = 0, \\ x, & \text{otherwise.} \end{cases}$$

Then $x \mapsto f``x$, aka $j(f)$, is a bijection from V to $\{x : \emptyset \notin x\}$. We call it '$\theta$' for short.

For any involution σ of any set X we define an embedding of involutions π from σ to $j(c)$ by

$$x \mapsto j(\theta)(x) \cup j(c \cdot \theta)(\sigma(x)).$$

https://doi.org/10.1017/jsl.2024.47 Published online by Cambridge University Press

The function π is injective, with left inverse

$$y \mapsto j(\theta^{-1})(\{z \in y : \emptyset \notin z\}).$$

To see that π is a map of involutions from σ to $j(c)$ we calculate as follows:

$$\begin{aligned}(j(c)\cdot\pi)(x) &= j(c)[j(\theta)(x) \cup j(c\cdot\theta)(\sigma(x))] \\ &=^{(1)} j(c\cdot\theta)(x) \cup j(c\cdot c\cdot\theta)(\sigma(x)) \\ &=^{(2)} j(\theta)(\sigma(x)) \cup j(c\cdot\theta)(\sigma(\sigma(x))) \\ &= (\pi\cdot\sigma)(x).\end{aligned}$$

(1) Distribute $j(c)$ over $\cup$;
(2) $c^2 = \mathbf{1}$ and reverse the order of the summands.

For the main result we argue as follows.

Clearly any involution into which a universal involution can be embedded is also universal, and any involution conjugate to a universal involution is again universal.

Since $j(c)$ is universal, there is an embedding of c into $j(c)$. This lifts to embeddings of $j^i(c)$ into $j^{i+1}(c)$, and composing these embeddings we get embeddings of $j(c)$ into $j^i(c)$ for any $i \geq 1$. Thus $j^i(c)$ is universal for any $i \geq 1$. ⊣

The following corollary of Lemma 1 is key.

Corollary 1. *Any two universal involutions are conjugate.*

Corollary 2. *Every model of NF has a permutation model with an internal $\in$-automorphism.*

Proof. It follows from Corollary 1 that $j(c)$ and $j^2(c)$ are conjugate, making $j(c)$ an example of a permutation which is conjugate to j of itself. It was shown in [1] that any model containing such a permutation π has a permutation model wherein π has become an (internal) $\in$-automorphism. ⊣

In [1] it is shown that there must be such a π, but that was on the assumption of AC_2, and of course we have here scrupulously eschewed AC_2.

It is a consequence of Corollary 2 that there can be no definable wellfounded extensional relation on the universe, since if there were we could prove by induction on it that the only $\in$-automorphism is the identity. In [2] we will draw the inference that NF is not synonymous with ZF or anything like it.

For the main result which follows later (Corollary 3) we will need involutions σ and τ such that there is a permutation π conjugating σ to $j(\tau)\cdot c$ and τ to $j(\sigma)\cdot c$. The next lemma exhibits such a pair of involutions, taking σ to be c_1 and τ to be c_2.

Lemma 3. *There is an involution that conjugates c with c_3 and commutes with c_2.*

Proof. We begin by choosing a fixed point a of c_2 and setting $b = c_1(a)$. (a could be $\{x : \emptyset \notin x\}$, but we don't need the extra detail: all we need is $a = c_2(a)$.) Since a is a fixed point of c_2 we also have $b = c_1(c_2(a)) = j(c)(a)$. For any $s \subseteq \{a, b\}$ we define X_s to be $\{x : x \cap \{a, b\} = s\}$. The X_s partition V into four pieces. $X_\emptyset$ is closed under both $j(c)$ and $j^2(c)$; let $\sigma_\emptyset$ be the restriction of $j(c)$ to $X_\emptyset$ and $\tau_\emptyset$ the restriction of $j^2(c)$.

https://doi.org/10.1017/jsl.2024.47 Published online by Cambridge University Press

Then there are embeddings of $j(c)$ into $\sigma_\emptyset$ and $j^2(c)$ into $\tau_\emptyset$, so by the results of the last section both $\sigma_\emptyset$ and $\tau_\emptyset$ are universal.

We had better justify this last paragraph. ⊣

LEMMA 4. *Let σ be any involution of V which doesn't send any natural number to a natural number, and let a and b be distinct sets swapped by σ. Let $X_\emptyset$ be the set of sets containing neither a nor b. Then there is an embedding of $j(\sigma)$ into its restriction to $X_\emptyset$.*

PROOF. Let n be a natural number such that, if either of a or b is a natural number, then it is less than n. We define $\pi\colon V \to V$ by

$$x \mapsto \begin{cases} n, & \text{if } x = a, \\ \sigma(n), & \text{if } x = b, \\ x+1, & \text{if } x \text{ is a natural number } \geq n, \\ \sigma(\sigma(x)+1), & \text{if } \sigma(x) \text{ is a natural number } \geq n, \\ x, & \text{otherwise.} \end{cases}$$

It is clear by construction that $\sigma \cdot \pi = \pi \cdot \sigma$, and that neither a nor b is in the image of π. But then also $j(\sigma) \cdot j(\pi) = j(\pi) \cdot j(\sigma)$ and the image of $j(\pi)$ is included in $X_\emptyset$, as required. ⊣

Let $\pi_\emptyset$ be an isomorphism from $\sigma_\emptyset$ to $\tau_\emptyset$. Since $j(c) = c_1 \cdot c_2$ and $j^2(c) = c_3 \cdot c_2$ we have the equation $\pi_\emptyset \cdot c_1 \cdot c_2 = c_3 \cdot c_2 \cdot \pi_\emptyset$, which we record for future use.

We now define $\pi\colon V \to V$ by

$$x \mapsto \begin{cases} \pi_\emptyset(x), & \text{if } x \cap \{a,b\} = \emptyset, \\ x, & \text{if } x \cap \{a,b\} = \{b\}, \\ c_3(c_1(x)), & \text{if } x \cap \{a,b\} = \{a\}, \\ c_3(\pi_\emptyset(c_1(x))), & \text{if } x \cap \{a,b\} = \{a,b\}. \end{cases}$$

The X_s form a partition of V, and π is a union of bijections from X_s to X_s for each $s \subseteq \{a,b\}$, so π is a permutation of V. It remains to verify that for any x we have both $\pi(c_1(x)) = c_3(\pi(x))$ and $\pi(c_2(x)) = c_2(\pi(x))$. For each equation there are four cases, depending on $x \cap \{a,b\}$.

We now check these cases for the first equation.

- If $x \cap \{a,b\} = \emptyset$, then $c_1(x) \cap \{a,b\} = \{a,b\}$ and so

$$\pi(c_1(x)) = c_3(\pi_\emptyset(c_1(c_1(x)))) = c_3(\pi_\emptyset(x)) = c_3(\pi(x)).$$

- If $x \cap \{a,b\} = \{b\}$ then $c_1(x) \cap \{a,b\} = \{a\}$ and so

$$\pi(c_1(x)) = c_3(c_1(c_1(x))) = c_3(x) = c_3(\pi(x)).$$

- If $x \cap \{a,b\} = \{a\}$ then $c_1(x) \cap \{a,b\} = \{b\}$ and so

$$\pi(c_1(x)) = c_1(x) = c_3(c_3(c_1(x))) = c_3(\pi(x)).$$

- If $x \cap \{a,b\} = \{a,b\}$ then $c_1(x) \cap \{a,b\} = \emptyset$ and so

$$\pi(c_1(x)) = \pi_\emptyset(c_1(x)) = c_3(c_3(\pi_\emptyset(c_1(x)))) = c_3(\pi(x)).$$

https://doi.org/10.1017/jsl.2024.47 Published online by Cambridge University Press

The four cases for the other equation are similar.

- If $x \cap \{a,b\} = \emptyset$ then $c_2(x) \cap \{a,b\} = \{a,b\}$ and so
$$\pi(c_2(x)) = c_3(\pi_\emptyset(c_1(c_2(x)))) = c_3(c_3(c_2(\pi_\emptyset(x)))) = c_2(\pi_\emptyset(x)) = c_2(\pi(x))\,.$$
- If $x \cap \{a,b\} = \{b\}$ then $c_2(x) \cap \{a,b\} = \{b\}$ and so
$$\pi(c_2(x)) = c_2(x) = c_2(\pi(x))\,.$$
- If $x \cap \{a,b\} = \{a\}$ then $c_2(x) \cap \{a,b\} = \{a\}$ and so
$$\pi(c_2(x)) = c_3(c_1(c_2(x)))) = c_2(c_3(c_1(x))) = c_2(\pi(x))\,.$$
- If $x \cap \{a,b\} = \{a,b\}$ then $c_2(x) \cap \{a,b\} = \emptyset$ and so
$$\pi(c_2(x)) = \pi_\emptyset(c_2(x)) = \pi_\emptyset(c_2(c_1(c_1(x)))) = c_2(c_3(\pi_\emptyset(c_1(x)))) = c_2(\pi(x))\,.$$

COROLLARY 3.
• *Every model of NF has a permutation model that contains two (internal) permutations σ and τ satisfying*
$(\forall xy)(x \in y \longleftrightarrow \sigma(x) \notin \tau(y))$ *and*
$(\forall xy)(x \in y \longleftrightarrow \tau(x) \notin \sigma(y))$.
• *Any such model satisfies duality for formulæ that are stratifiable-mod-2.*

PROOF. We use the permutation π from Lemma 3, and exploit the two permutations σ and τ that we find in the permutation model V^π.

If a formula ϕ is stratifiable-mod-2 then its variables can be assigned to two types `yin` and `yang` in such a way that in subformulæ like '$x = y$' the two variables receive the same type and in subformulæ like '$x \in y$' the two variables receive different types. Let us associate σ to variables given type `yin` in the assignment and associate τ to variables given type `yang` in the assignment. '$x \in y$' is equivalent to '$\sigma(x) \notin \tau(y)$' and if x is of type `yin` we make this replacement. '$x \in y$' is also equivalent to '$\tau(x) \notin \sigma(y)$' and if x is of type `yang` we make this replacement. We deal with equations analogously. In the rewritten version of ϕ we find that every variable 'x' of type `yin` now appears only as '$\sigma(x)$' and that every variable 'y' of type `yang` now appears only as '$\tau(y)$'. So we can reletter '$\sigma(x)$' as 'x', and '$\tau(y)$' as 'y' and the result is $\phi^{\circlearrowleft}$. ⊣

We do not believe that Corollary 3 is best possible.

REFERENCES

[1] T. FORSTER, ***Set Theory with a Universal Set***, Oxford Logic Guides, 20, Oxford University Press, Oxford, 1992.

[2] T. FORSTER and R. HOLMES, *Synonymy questions concerning the Quine systems*, ***Journal of Symbolic Logic***, to appear.

FACHBEREICH MATHEMATIK
UNIVERSITÄT HAMBURG
BUNDESSTRAßE 55
20146 HAMBURG
GERMANY
E-mail: nathan.bowler@uni-hamburg.de

SCHOOL OF MATHEMATICS AND STATISTICS
VICTORIA UNIVERSITY OF WELLINGTON
WELLINGTON 6012
NEW ZEALAND
E-mail: tf@dpmms.cam.ac.uk