

Permutation Models for Set Theory

Permutation Models for Set Theory

Contents

1	Introduction	1
1.1	Preliminaries	2
1.2	Acknowledgements	5
2	The Boolean prime ideal theorem	6
2.1	Some equivalences	9
3	Set theory with atoms	16
3.1	Fraenkel–Mostowski permutation models	18
3.2	Parameter modification	19
3.3	Normal filter–normal ideal duality	21
4	Mostowski’s ordered model	25
5	Topological groups and their actions	28
5.1	Extremely amenable groups	28
5.2	Ramsey filters	31
6	The correspondence	33
6.1	Post mortem	38
	References	40

1 Introduction

Permutation models of set theory were devised by Abraham Fraenkel in the 1920s to prove the independence of the axiom of choice from other axioms. They were found very useful in the subsequent decades for establishing non-implications—for instance, that the ordering principle does not imply the axiom of choice—but soon fell out of use for this purpose when Paul Cohen introduced the forcing method in the 1960s. The reason for this is that permutation models require a weakening of the axiom of extensionality to allow for the existence of *atoms* (empty objects which are not sets; see Section 3), whereas forcing allows one to work entirely within the more conventional set theory ZFC.

Nonetheless, the methods used in the study of permutation models live on. Permutation methods underlie forcing via Boolean-valued models, nominal sets in computer science (e.g. [18]), and a proposed proof due to Randall Holmes that Quine’s New Foundations set theory is consistent relative to ZF, a question which has been open for seventy-six years.

The focus of this essay is Andreas Blass’s recent result [4] that permutation models in which the Boolean prime ideal theorem holds but the axiom of choice fails correspond with nontrivial extremely amenable topological groups with small open subgroups. The reader is not expected to understand the statement of this result at this stage; the necessary theory will be developed throughout the essay.

Blass’s result is surprising for a number of reasons. Firstly, it draws from a diverse range of distinct fields of mathematics, most notably set theory, group theory, combinatorics and topology. Secondly, the theory of amenable groups, more specifically of extremely amenable groups, was not contrived in any way to prove results in set theory: the question of the existence of nontrivial extremely amenable groups was asked by Theodore Mitchell [16] in a paper about topological semigroups and answered by Wojciech Herer and Jens Christensen [8] five years later in a measure theory paper. Further examples were given by Vladimir Pestov [17] in a paper about topological dynamics.

What I hope to achieve with this essay is to give a self-contained exposition of Blass’s theorem, proving (almost) all results in full and building up all the necessary grounding. In Section 2 the Boolean prime ideal theorem is explored in depth: we establish how strong a statement it is relative to other weakenings of the axiom of choice and prove that it is equivalent to a number of other assertions in mathematics. Set theory with atoms and Fraenkel–Mostowski permutation models are introduced in Section 3 and their basic theory explored, including a digression to discuss normal filters and normal ideals in greater depth. Section 4 discusses a particular example

of a permutation model, namely Mostowski's ordered model, which will serve to illustrate the main theorem of the essay. After a digression in Section 5 to build up the group theory and topology needed to understand the other half of the correspondence, we will be ready to state the result in full detail and prove it in Section 6.

1.1 Preliminaries

This part of the essay is devoted to briskly introducing some concepts which underpin the rest of the essay—particularly those outside of set theory. A basic grounding of sentential logic, first-order logic and axiomatic set theory is assumed, such as that found in [6] or [12], as is an elementary knowledge of group theory and general topology.

Partially ordered sets

A *partially ordered set* $\mathbb{P} = \langle P, \leq \rangle$ is a set P together with a reflexive, transitive, antisymmetric binary relation \leq . I will abuse notation slightly by writing $x \in \mathbb{P}$ to mean $x \in P$, and omit reference to the underlying set. The *strict part* of the order relation \leq is denoted by $<$. The *discrete order* is defined by $p \leq q \rightarrow p = q$ for all $p, q \in \mathbb{P}$.

\mathbb{P} is *totally ordered* if for each $x, y \in \mathbb{P}$ either $x \leq y$ or $y \leq x$. The *order extension principle* is the assertion that every partial order can be extended to a total order. The *ordering principle* is the assertion that every set can be endowed with a total order. A subset $C \subseteq \mathbb{P}$ is a *chain* if it is totally ordered by the restriction of \leq to C .

\mathbb{P} is *wellordered* if it is totally ordered and every nonempty subset of \mathbb{P} has a $<$ -least member. The *wellordering principle* is the assertion that every set can be wellordered, or equivalently that for every set x there is an ordinal α_x and an injective function $x \rightarrow \alpha_x$.

A function $f : \mathbb{P} \rightarrow \mathbb{P}'$ is *order-preserving* if for all $x, y \in \mathbb{P}$, $x \leq y$ implies $f(x) \leq f(y)$, *order-reversing* if $x \leq y$ implies $f(y) \leq f(x)$ and *monotonic* if it is order-preserving or order-reversing. It is an *order-isomorphism* if it is bijective and both f and f^{-1} are order-preserving, in which case we write $\mathbb{P} \cong \mathbb{P}'$. An *order-automorphism* of \mathbb{P} is an order-isomorphism $\mathbb{P} \rightarrow \mathbb{P}$.

A *lattice* \mathbb{L} is a partially ordered set with all finite *joins* (suprema, \vee) and *meets* (infima, \wedge). In particular, every lattice has a least element $0 = \bigvee \emptyset$ and a greatest element $1 = \bigwedge \emptyset$. A lattice is *complete* if it has all joins. In this essay, all lattices will be assumed to be *nontrivial*, i.e. $0 \neq 1$. For a pair $x, y \in \mathbb{L}$ write $x \vee y = \bigvee \{x, y\}$ and $x \wedge y = \bigwedge \{x, y\}$.

A lattice \mathbb{L} is *distributive* if for any $x, y, z \in \mathbb{L}$ we have

$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z) \quad \text{and} \quad x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$$

Lemma 1.1 *If \mathbb{L} is a distributive lattice then there exists a set X and a lattice $\tilde{\mathbb{L}}$ of subsets of X such that $\mathbb{L} \cong \tilde{\mathbb{L}}$. \square*

A distributive lattice \mathbb{B} is a *Boolean algebra* if every element $x \in \mathbb{B}$ has a *complement* $\neg x \in \mathbb{B}$, which is unique with the properties

$$x \wedge \neg x = 0, \quad x \vee \neg x = 1, \quad \neg \neg x = x$$

The *dual algebra* \mathbb{B}^{op} of a Boolean algebra \mathbb{B} has the same underlying set but with reversed order.

The axiom of choice and Zermelo–Fraenkel set theory

The *axiom of choice* (AC) is the assertion that every set of nonempty sets has a *choice function*; that is, if I is a set and $\{x_\alpha : \alpha \in I\}$ is a set of nonempty sets then there is a function $f : I \rightarrow \bigcup_{\alpha \in I} x_\alpha$ such that $f(\alpha) \in x_\alpha$ for each $\alpha \in I$.

Zorn’s lemma is the assertion that if \mathbb{P} is a partially ordered set all of whose chains have an upper bound, then \mathbb{P} has a maximal element.

Lemma 1.2 *The axiom of choice is equivalent to both the wellordering principle and Zorn’s lemma. \square*

ZF denotes Zermelo–Fraenkel set theory, ZFC denotes ZF with AC appended as an axiom, and ZFA denotes a modification of ZF to be defined in Section 3. We will work inside ZF(A) wherever possible; choice is assumed to hold in the ‘real world’ (the metatheory) and will be mentioned explicitly wherever used.

A formula ϕ is *independent* of a theory T if $\mathsf{T} \not\vdash \phi$ and $\mathsf{T} \not\vdash \neg\phi$, and is independent of another formula ψ *in the presence* of T if $\mathsf{T} \not\vdash \psi \rightarrow \phi$. Similarly, ϕ and ψ are *equivalent* in T if $\mathsf{T} \vdash \phi \leftrightarrow \psi$. I omit reference to T when the independence or equivalence holds in each of ZF, ZFC and ZFA.

$\mathcal{P}(x)$ denotes the *powerset* of a set x . The *transitive closure* of a set x is defined by $\text{trcl}(x) = \bigcup \{x, \bigcup x, \bigcup \bigcup x, \dots\}$. If ϕ is a predicate with one free variable then $\text{H}(\phi)$ is the class of *hereditarily* ϕ sets, defined by

$$x \in \text{H}(\phi) \leftrightarrow \phi(x) \wedge \forall y (y \in \text{trcl}(x) \rightarrow \phi(y))$$

The symbols \mathbb{N} , \mathbb{Z} , \mathbb{Q} , \mathbb{R} denote the sets of *natural numbers*, *integers*, *rational numbers* and *real numbers*, respectively; we take 0 to be a natural number. Ordinals are taken to be *von Neumann ordinals*, i.e. sets α which are wellordered by \in and such that $\beta \in \alpha \rightarrow \beta \subseteq \alpha$. The *successor* of an ordinal α is $\alpha + 1 = \alpha \cup \{\alpha\}$; ordinals not of this form are called *limit ordinals*. The class of all ordinals is denoted by **Ord**. We adopt the usual habit of identifying finite ordinals with natural numbers.

I will use arrow notation f^\rightarrow , f^\leftarrow for the *direct image* and *inverse image* of a function f . Specifically, if $f : X \rightarrow Y$ is a function, $A \subseteq X$ and $B \subseteq Y$, then

$$f^\rightarrow(A) = \{f(x) : x \in A\} \quad \text{and} \quad f^\leftarrow(B) = \{x \in X : f(x) \in B\}$$

It is important to distinguish between $f(A)$ and $f^\rightarrow(A)$ whenever $A \cup \{A\} \subseteq X$, especially if f is defined on the entire set-theoretic universe, because $f(A)$ and $f^\rightarrow(A)$ may not be equal.

General topology

A (*topological*) *space* is a set X equipped with a *topology* $\mathcal{O}(X)$, which is a lattice of subsets of X closed under arbitrary joins (unions) and finite meets (intersections). Elements of $\mathcal{O}(X)$ are called *open sets* and their complements in X are called *closed sets*. If $\mathcal{O}(X)$ is a topology on X and $\mathcal{B} \subseteq \mathcal{O}(X)$ then $\mathcal{O}(X)$ is *generated by* \mathcal{B} , and \mathcal{B} is a *base* for $\mathcal{O}(X)$, if $\mathcal{O}(X)$ is the closure of \mathcal{B} under arbitrary unions.

The *discrete topology* on a space X is $\mathcal{P}(X)$. If $\{X_i : i \in I\}$ is a set of topological spaces and $X = \prod_{i \in I} X_i = \{(f : I \rightarrow X_i) : f(i) \in X_i\}$ then define $\pi_i : X \rightarrow X_i$ by $\pi_i(f) = f(i)$; the *product topology* on X is the topology generated by sets of the form $\pi_i^{-1}(U)$ for U open in X_i .

A space X is *Hausdorff* if for any distinct points $x, y \in X$ there are disjoint open sets U, V with $x \in U$ and $y \in V$. X is *compact* if whenever $\{U_\alpha : \alpha \in I\}$ is a set of open sets with $\bigcup_{\alpha \in I} U_\alpha = X$, there is a finite subset $J \subseteq I$ such that $\bigcup_{\alpha \in J} U_\alpha = X$.

Lemma 1.3 *If X is a compact Hausdorff space and $x, y \in X$ are distinct then there exists a continuous map $f : X \rightarrow [0, 1]$ such that $f(x) = 0$ and $f(y) = 1$.* \square

Groups and group actions

If G is a group then $\text{Sub}(G)$ denotes the lattice of subgroups of G , where $H \wedge K = H \cap K$ and $H \vee K$ is the subgroup generated by the elements of

$H \cup K$. $H \leq G$ denotes that H is a subgroup of G , and $H \trianglelefteq G$ denotes that H is a normal subgroup of G .

If X is a set, then a *left action* of G on X will be denoted by $G \curvearrowright X$. This will be formalised as a map $G \times X \rightarrow X$ given by $\langle g, x \rangle \mapsto g \cdot x$, which satisfies $1 \cdot x = x$ and $(gh) \cdot x = g \cdot (h \cdot x)$ for all $g, h \in G$ and $x \in X$. The *orbit* of an element $x \in X$ will be denoted by

$$G(x) = \{g \cdot x : g \in G\}$$

1.2 Acknowledgements

First and foremost, I am extremely grateful to Thomas Forster. He is responsible for nurturing my interest in logic and set theory, and for sparking my interest in permutation models through the Fraenkel–Mostowski permutation models reading group which he organised in Michaelmas Term 2012. I have enjoyed indulging in numerous discussions with him about logic and set theory, and I owe him a great debt for his undying support over the last two years.

I would also like to thank Andreas Blass for being kind enough to point me in the direction of his paper [4], from which the most important result of the essay is drawn, and for responding to my questions [5].

Finally I would like to thank the members of the Fraenkel–Mostowski models reading group and the $\text{con}(\text{NF})$ discussion group, especially Philipp Kleppmann, who provided me with a paper [15] on which Section 3.3 is based.

A complete list of papers that I have read and from which ideas have been drawn can be found in the list of references on page 41. I have drawn most heavily from Thomas Jech’s book *The Axiom of Choice* [10] and Andreas Blass’s papers [3] and [4], from which I learnt the material on which the essay is based. Many of the proofs in the essay are adapted directly from these references. The results whose proofs are my own are Lemmas 2.1, 4.1 and 6.6; Propositions 2.3, 2.5, 2.13, 3.1, 3.6, 3.7*, 5.1* and 5.2; and Theorem 6.7. Those marked with * are my own results.

My construction of ZFA adopts the convention that the collection of atoms needn’t form a set, even though for the purposes of this essay it will always be a set. My definitions of *mess* and *consistency principle* (page 10) generalise those usually found in the literature (e.g. [10]); Proposition 2.6 is generalised accordingly.

2 The Boolean prime ideal theorem

Let \mathbb{L} be a distributive lattice and let $I \subseteq \mathbb{L}$ be a nonempty subset. Consider the following properties that I may, or may not, have:

- (a) I is closed downwards: if $y \in I$ and $x \leq y$ then $x \in I$;
- (b) I is closed under pairwise joins: if $x, y \in I$ then $x \vee y \in I$;
- (c) I is proper: $I \neq \mathbb{L}$;
- (d) $\mathbb{L} - I$ is closed under joins: if $x \wedge y \in I$ then $x \in I$ or $y \in I$.
- (e) I is maximal: if $I \subseteq J$ and J satisfies (a)–(c) then $I = J$.

Then I is:

- an *ideal* if it satisfies (a)–(b);
- a *proper ideal* if it satisfies (a)–(c);
- a *prime ideal* if it satisfies (a)–(d);
- a *maximal ideal* if it satisfies (a)–(c) and (e).

Dually, there are notions of *filter*, *proper filter*, *prime filter* and *ultrafilter*, which are obtained by swapping \vee and \wedge , and replacing \leq by \geq , in (a)–(e) above. For example, a filter is an upwards-closed set which is closed under pairwise meets. Note that, for an ideal I , (c) is equivalent to the assertion that $1 \notin I$.

For a Boolean algebra \mathbb{B} , there is another property that I may have:

- (*) For all $b \in \mathbb{B}$, either $b \in I$ or $\neg b \in I$, but not both.

Lemma 2.1 *Let \mathbb{B} be a Boolean algebra. A proper ideal $I \subseteq \mathbb{B}$ is prime if and only if it is maximal. Furthermore, this holds if and only if (*) holds.*

Proof We prove that primality (d) and maximality (e) are both equivalent to (*) in the presence of properness (a)–(c).

(d) \leftrightarrow (*). If I is prime and $b \in \mathbb{B}$ then $b \wedge \neg b = 0 \in I$ and so $b \in I$ or $\neg b \in I$ by (d), so (*) holds. Conversely, if (*) holds and $a, b \in \mathbb{B}$ with $a \notin I$ and $b \notin I$ then $\neg a \in I$ and $\neg b \in I$. Then $\neg a \vee \neg b = \neg(a \wedge b) \in I$ by (b), so $a \wedge b \notin I$ by (c), and so I is prime.

(e) \leftrightarrow (*). Let I be a proper ideal and suppose $b \in \mathbb{B}$ is a witness to the failure of (*), so that $b \notin I$ and $\neg b \notin I$. We must have $b < 1$ because $\neg 1 = 0 \in I$. Let

$$I^+ = I \cup \{a \in \mathbb{B} : a \leq b\} \cup \{a \vee b : a \in I\}$$

It is plain to see that I^+ is closed downwards and under joins. The only ways in which it could fail to be proper are if we had either $1 \leq b$ and hence $b = 1$, which is impossible, or $1 = a \vee b$ for some $a \in I$, in which case

$$a = a \vee 0 = a \vee (b \wedge \neg b) = (a \vee b) \wedge (a \vee \neg b) = 1 \wedge (a \vee \neg b) = a \vee \neg b$$

But then $\neg b \leq a$, and hence $\neg b \in I$, contradicting our assumption. So I^+ is a proper ideal strictly containing I , and I is not maximal. Conversely, suppose (*) holds and that J is a proper ideal with $I \subseteq J$. If $b \in J$ but $b \notin I$ then $\neg b \in I$, so $\neg b \in J$ and hence $1 = b \vee \neg b \in J$ contradicting properness of J ; hence $I = J$ and so I is maximal.

Hence (d) \leftrightarrow (e) \leftrightarrow (*). And if any of these hold and $b \in I$ then $\neg b \notin I$, else $b \vee \neg b = 1 \in I$, contradicting maximality. \square

The *Boolean prime ideal theorem* (BPIT) is the assertion that any proper ideal in a Boolean algebra can be extended to a prime ideal. In fact this statement is equivalent to an apparently weaker one:

Lemma 2.2 *The Boolean prime ideal theorem is equivalent to the assertion that every Boolean algebra has a prime ideal.*

Proof The latter statement is implied by BPIT since $\{0\}$ is an ideal in any Boolean algebra B and any ideal must contain 0.

Conversely, suppose every Boolean algebra has a prime ideal and let I be an ideal in a Boolean algebra \mathbb{B} . Define a relation \sim on \mathbb{B} by

$$a \sim b \quad \text{if and only if} \quad (a \wedge \neg b) \vee (\neg a \wedge b) \in I$$

Then \sim is an equivalence relation. It is reflexive since $0 \in I$, it is clearly symmetric, and it is transitive since if $a \sim b$ and $b \sim c$ then by closure under unions we have

$$(a \wedge \neg b) \vee (\neg a \wedge b) \vee (b \wedge \neg c) \vee (\neg b \wedge c) \in I$$

and so we obtain

$$\begin{aligned}
(a \wedge \neg c) \vee (\neg a \wedge c) &= (a \wedge b \wedge \neg c) \vee (a \wedge \neg b \wedge \neg c) \\
&\quad \vee (\neg a \wedge b \wedge c) \vee (\neg a \wedge \neg b \wedge c) \\
&\leq (a \wedge b \wedge \neg c) \vee (a \wedge \neg b \wedge \neg c) \vee (a \wedge \neg b \wedge c) \\
&\quad \vee (\neg a \wedge b \wedge c) \vee (\neg a \wedge \neg b \wedge c) \vee (\neg a \wedge b \wedge \neg c) \\
&= (a \wedge \neg b) \wedge (c \vee \neg c) \vee (\neg a \wedge b) \wedge (c \vee \neg c) \\
&\quad \vee (b \wedge \neg c) \wedge (a \vee \neg a) \vee (\neg b \wedge c) \wedge (a \vee \neg a) \\
&= (a \wedge \neg b) \vee (\neg a \wedge b) \vee (b \wedge \neg c) \vee (\neg b \wedge c) \in I
\end{aligned}$$

Hence $(a \wedge \neg c) \vee (\neg a \wedge c) \in I$ by downwards-closure, and so $a \sim c$. So \sim is transitive, and hence an equivalence relation.

Define operations on the set \mathbb{B}/\sim of \sim -equivalence classes pointwise; then \mathbb{B}/\sim is a Boolean algebra, and hence contains a prime ideal Q say. But then

$$P = \{b \in \mathbb{B} : [b]_{\sim} \in Q\}$$

is a prime ideal in \mathbb{B} containing I . □

Proposition 2.3 *The axiom of choice implies the Boolean prime ideal theorem.*

Proof Suppose the axiom of choice holds and let \mathbb{B} be a Boolean algebra.

The partially ordered set of proper ideals in \mathbb{B} is chain-complete; this can be easily verified by noting that the supremum of a chain is simply its union, which is proper. Let I be any proper ideal in \mathbb{B} . By Zorn's lemma, which is equivalent to the axiom of choice by Lemma 1.2, there is a proper ideal P containing I which is \subseteq -maximal amongst all proper ideals of \mathbb{B} ; i.e. it is a maximal ideal, and hence is prime by Lemma 2.1. □

In fact, there is a chain of implications

$$\text{axiom of choice} \rightarrow \text{BPIT} \rightarrow \text{order-extension principle} \rightarrow \text{ordering principle}$$

The first implication is Proposition 2.3, the second follows immediately from Theorem 2.4 and Lemma 2.7 below, and the third is an application of the order extension principle to sets with the discrete ordering. We will see in Section 4 that in Mostowski's 'ordered' model of ZFA the ordering principle holds but the axiom of choice does not. As such, Mostowski's ordered model must exhibit the failure of one of these reverse implications. It is striking that the axiom of choice fails so weakly in this model that BPIT still holds, thus proving that BPIT is independent from the axiom of choice in the presence of the ZFA axioms.

2.1 Some equivalences

In order to prove this independence result and the main result of the paper, we must venture through numerous areas of mathematics and logic, each of which lends itself to different ways of thinking. The Boolean prime ideal theorem, as stated, is not much use for many of our purposes. As such, the remainder of this section is devoted to proving that BPIT is equivalent to a number of different statements. Namely, we introduce the *consistency principle* (CP), the *ultrafilter lemma* (UL), the *sentential compactness theorem* (SCT) and the *almost maximal ideal theorem* (AMIT), and prove results working up to the following theorem.

Theorem 2.4 *The following are equivalent:*

- *Boolean prime ideal theorem*
- *Ultrafilter lemma*
- *Consistency principle*
- *Sentential compactness theorem*
- *Almost maximal ideal theorem*

Proof The propositions in the remainder of Section 2 give the following chain of implications:

$$\text{BPIT} \xrightarrow{(2.5)} \text{UL} \xrightarrow{(2.6)} \text{CP} \xrightarrow{(2.8)} \text{SCT} \xrightarrow{(2.12)} \text{AMIT} \xrightarrow{(2.13)} \text{BPIT}$$

so the equivalences are established. \square

For example, to prove $\text{ZFA} \not\vdash \text{BPIT} \rightarrow \text{AC}$ we will instead prove $\text{ZFA} \not\vdash \text{CP} \rightarrow \text{AC}$ (Theorem 4.7), and to prove, for a particular model \mathbb{M} of ZFA, that $\mathbb{M} \models \text{BPIT}$ is equivalent to a certain condition we shall prove that the condition is implied by SCT and implies AMIT (Theorem 6.5).

Ultrafilter lemma

Ultrafilters are dual to maximal ideals in the sense that U is an ultrafilter in a Boolean algebra \mathbb{B} if and only if it is a maximal ideal in the dual algebra \mathbb{B}^{op} .

The *ultrafilter lemma* (UL) is the assertion that any proper filter on a Boolean algebra can be extended to an ultrafilter.

Proposition 2.5 *The Boolean prime ideal theorem implies the ultrafilter lemma.*

Proof Let \mathbb{B} be a Boolean algebra and $F \subseteq \mathbb{B}$ be a proper filter. Then F is a proper ideal in \mathbb{B}^{op} , so it can be extended to a prime ideal U in \mathbb{B}^{op} . By Lemma 2.1, U is a maximal ideal in \mathbb{B}^{op} , and so U is an ultrafilter in \mathbb{B} . \square

Consistency principle

We will now reformulate BPIT in combinatorial terms. A *mess* on a lattice \mathbb{L} is a pair $\mathcal{M} = \langle M, I \rangle$, where I is an ideal in \mathbb{L} and M is a set of pairs $\langle x, y \rangle$ with $x \in I$ and $y \leq x$, such that:

- For all $x \in I$ there exists $y \leq x$ with $\langle x, y \rangle \in M$.
- If $\langle x, y \rangle \in M$ and $z \in I$ then $\langle z, y \wedge z \rangle \in M$.

We say $\langle x, y \rangle$ and $\langle x', y' \rangle$ are *compatible* if $x \wedge y' = x' \wedge y$. If $c \in \mathbb{L}$ is an arbitrary element and \mathcal{M} is a mess on \mathbb{L} then c is *consistent with* \mathcal{M} if $\langle x, c \wedge x \rangle \in M$ for each $x \in I$. When $\mathbb{L} = \mathcal{P}(X)$ for some set X and I is the ideal of finite subsets of X , we identify \mathcal{M} with M and say \mathcal{M} is a *mess on* X .

A complete lattice \mathbb{F} is called a *frame* if for all sets I and all $x, y_i \in \mathcal{F}$ ($i \in I$)

$$x \wedge \bigvee_{i \in I} y_i = \bigvee_{i \in I} x \wedge y_i \quad \text{and} \quad x \vee \bigwedge_{i \in I} y_i = \bigwedge_{i \in I} x \vee y_i$$

Note in particular that all power sets are frames.

The *consistency principle* (CP) is the assertion that if \mathbb{F} is a frame and $\mathcal{M} = \langle M, I \rangle$ is a mess on \mathbb{F} , then there is an element $c \in \mathbb{F}$ which is consistent with \mathcal{M} .

Proposition 2.6 *The ultrafilter lemma implies the consistency principle.*

Proof Let $\mathcal{M} = \langle M, I \rangle$ be a mess on a frame \mathbb{F} . Let A be the set of all functions $j : J \rightarrow I$ with $J \subseteq I$ such that

- (a) $j(x) \leq x$ and $\langle x, j(x) \rangle \in M$ for all $x \in J$;
- (b) If $x, y \in J$ then $\langle x, j(x) \rangle$ and $\langle y, j(y) \rangle$ are compatible.

For each $x \in I$ and $y \leq x$, define

$$A_x = \{(j : J \rightarrow I) \in A : x \in J\}$$

$$B_{x,y} = \{(j : J \rightarrow I) \in A : x \in J \text{ and } j(x) = y\}$$

Let \mathcal{F} be the filter generated by $\{A_x : x \in I\}$. Each A_x is nonempty: if $x \in I$ then there exists $y \leq x$ with $\langle x, y \rangle \in M$; then the function $j : \{x\} \rightarrow I$ given by $j(x) = y$ lies in A_x . Furthermore, if $x, x' \in I$ then $A_x \cap A_{x'} \neq \emptyset$. To see this, choose y, y' such that $\langle x, y \rangle \in M$ and $\langle x', y' \rangle \in M$. By property (b) of messes, $\langle y', y \wedge y' \rangle \in M$ and so $\langle x, x \wedge y \wedge y' \rangle = \langle x, y \wedge y' \rangle \in M$, and likewise $\langle x', y \wedge y' \rangle \in M$. Define $j : \{x, x'\} \rightarrow I$ by $j(x) = j(x') = y \wedge y'$. Then

$$x' \wedge j(x) = x' \wedge y \wedge y' = y \wedge y' = x \wedge y \wedge y' = x \wedge j(x')$$

So $\langle x, j(x) \rangle$ and $\langle x', j(x') \rangle$ are compatible, and $j \in A_x \cap A_{x'}$. Hence \mathcal{F} is a proper filter.

By the ultrafilter lemma, there is an ultrafilter \mathcal{U} on A extending \mathcal{F} . Now, for each $x \in I$, we can write A_x as a disjoint union:

$$A_x = B_{x, y_1} \cup \dots \cup B_{x, y_m}$$

where $\langle x, y_1 \rangle, \dots, \langle x, y_m \rangle$ exhausts all elements of M whose first component is x . Hence there exists $x_* \leq x$ with $\langle x, x_* \rangle \in M$ such that $B_{x, x_*} \in \mathcal{U}$. This x_* is unique: the B_{x, y_i} are disjoint, so if more than one were to lie in \mathcal{U} then so would their intersection, and then we'd have $\emptyset \in \mathcal{U}$, contradicting properness. Then for any $x, z \in I$, $B_{x, x_*} \cap B_{z, z_*} \neq \emptyset$ by properness of \mathcal{U} . By definition of A , $\langle x, x_* \rangle$ and $\langle z, z_* \rangle$ are compatible, so $z \wedge x_* = x \wedge z_*$.

Define $c = \bigvee_{x \in I} x_*$. Then for any $z \in I$,

$$z \wedge c = \bigvee_{x \in I} z \wedge x_* = \bigvee_{x \in I} x \wedge z_* = z_*$$

Hence $\langle z, z \wedge c \rangle \in M$, and c is consistent with \mathcal{M} . \square

Sentential compactness theorem

Our viewpoint now turns to that of logic; in particular, sentential logic. Let P be a set of sentential variables and Σ be a set of sentences over P . The *sentential compactness theorem* (SCT) states that if each finite subset $\Sigma_0 \subseteq \Sigma$ has a model then Σ has a model.

Lemma 2.7 *The sentential compactness theorem implies the order extension principle.*

Proof Let $\mathbb{P} = \langle P, \leq \rangle$ be a partially ordered set. Define sentential variables $\llbracket p \leq^t q \rrbracket$ for $p, q \in P$ encoded, say, as elements of $P \times P$. Let

$$\begin{aligned} \Sigma = \{ & \llbracket p \leq^t q \rrbracket \wedge \llbracket q \leq^t r \rrbracket \Rightarrow \llbracket p \leq^t r \rrbracket : p, q, r \in P \} && \leftarrow \text{transitivity} \\ & \cup \{ \llbracket p \leq^t q \rrbracket \wedge \llbracket q \leq^t p \rrbracket \Rightarrow \llbracket p = q \rrbracket : p, q \in P \} && \leftarrow \text{symmetry} \\ & \cup \{ \llbracket p \leq^t q \rrbracket \vee \llbracket q \leq^t p \rrbracket : p, q \in P \} && \leftarrow \text{totality} \\ & \cup \{ \llbracket p \leq^t q \rrbracket : p, q \in P \text{ such that } p \leq q \} && \leftarrow \text{extension} \end{aligned}$$

Let $\Sigma_0 \subseteq \Sigma$ be finite, and let $P_0 \subseteq P$ be the finite subset of P consisting of all elements appearing in sentences in Σ_0 . It follows from the finite axiom of choice that P_0 has a total ordering extending the restriction to P_0 of \leq , and hence Σ_0 has a model. By the sentential compactness theorem, Σ has a model, which is to say that P has a total ordering \leq^t extending \leq . \square

Proposition 2.8 *The consistency principle implies the sentential compactness theorem.*

Proof Let P be a set of sentential variables and Σ be a set of sentences over P . Suppose each $\Sigma_0 \subseteq \Sigma$ has a model $v_{\Sigma_0} : P \rightarrow \{\top, \perp\}$. Let $P_{\Sigma_0} \subseteq P$ be the set of sentential variables appearing in formulae in Σ_0 . For $F \subseteq P_{\Sigma_0}$ let $F^\top = \{p \in F : v_{\Sigma_0}(p) = \top\}$. Define

$$\mathcal{M} = \{\langle F, E \rangle : \Sigma_0 \subseteq \Sigma \text{ finite, } F \subseteq P_{\Sigma_0}, E \subseteq F^\top\}$$

Each P_{Σ_0} must be finite since Σ_0 is finite and each formula in Σ_0 contains only finitely many variables, so each F and E are certainly finite. Furthermore, for each $F \subseteq P_{\Sigma_0}$, $\langle F, F^\top \rangle \in \mathcal{M}$; and if $\langle F, E \rangle \in \mathcal{M}$ and $G \subseteq P_{\Sigma_0}$ then $E \cap G \subseteq G^\top$ and so $\langle G, E \cap G \rangle \in \mathcal{M}$. Hence \mathcal{M} is a mess on P , and so there is some subset $C \subseteq P$ which is consistent with \mathcal{M} . Define $v : P \rightarrow \{\top, \perp\}$ by

$$v(p) = \begin{cases} \top & \text{if } p \in C \\ \perp & \text{if } p \notin C \end{cases}$$

Then for each finite $\Sigma_0 \subseteq \Sigma$ and $F \subseteq P_{\Sigma_0}$, $\langle F, C \cap F \rangle \in \mathcal{M}$, so $C \cap F \subseteq F^\top$. Hence v is a model of Σ . \square

Almost maximal ideal theorem

Let \mathbb{L} be a nontrivial distributive lattice (with 0 and 1) and let $I \subseteq \mathbb{L}$ be an ideal. The *dual* I^* of I is given by

$$I^* = \{v \in \mathbb{L} : u \vee v = 1 \text{ for some } u \in I\}$$

Lemma 2.9 *I is proper if and only if $I \cap I^* = \emptyset$*

Proof If $u \in I \cap I^*$ then $u \vee v = 1$ for some $v \in I$ and hence $1 \in I$, so I is not proper. Conversely if $1 \in I$ then $I \subseteq I^*$ since $1 \vee v = 1$ and so $I \cap I^* \neq \emptyset$. \square

We say that I is an *almost maximal ideal* if it is prime and its dual is maximal, in the sense that if $J \subseteq I$ and $J^* = I^*$ then $J = I$. More concisely, for an ideal $I \subseteq \mathbb{L}$ define

$$\widehat{I} = \{u \in \mathbb{L} : v \in I^* \text{ whenever } u \vee v = 1\}$$

Then I is almost maximal if it is prime and $I = \widehat{I}$.

The *almost maximal ideal theorem* (AMIT) states that every nontrivial distributive lattice has an almost maximal ideal.

To prove that AMIT is equivalent to BPIT we first need to build up some machinery. We will use the sentential compactness theorem along the way.

First we fix some notation. Let \mathbb{L} be a nontrivial distributive lattice with 0 and 1. Let P be the set of sentential variables $\llbracket x \vee y \in I \rrbracket$, where $I \subseteq \mathbb{L}$ is a proper ideal and $x, y \in \mathbb{L}$ with $x \vee y \in I$. This can be formalised by encoding $\llbracket x \vee y \in I \rrbracket$ as $\langle I, x, y \rangle$, for example. We define a set Σ of sentences over P as follows. Let X be the set of all triples $\langle I, F, f \rangle$ where

- $I \subseteq \mathbb{L}$ is a proper ideal;
- $F \subseteq \mathbb{L}$ is a finite subset with $\bigvee F \in I^*$;
- $f : F \rightarrow \mathbb{L}$ is a function with $x \wedge f(x) \in I$ for each $x \in F$.

For $\langle I, F, f \rangle \in X$ define

$$\xi(I, F, f) = \bigvee_{x \in F} \neg \llbracket x \vee f(x) \in I \rrbracket, \quad \eta(I, F, f) = \bigvee_{x \in F} \llbracket f(x) \vee x \in I \rrbracket$$

and let

$$\Sigma = \{\xi(I, F, f) : \langle I, F, f \rangle \in X\} \cup \{\eta(I, F, f) : \langle I, F, f \rangle \in X\}$$

Lemma 2.10 *The sentential compactness theorem implies that Σ is consistent.*

Proof In any sentence in Σ , each variable refers to just one ideal. Let $\Sigma_0 \subseteq \Sigma$ be finite, and assume that the only variables that appear in sentences in Σ_0 refer to the ideal I . If each Σ_0 of this form has a model then so do all finite subsets of Σ since any finite subset will be a finite union of (finite) sets of this form.

Let $Q = \{\langle x, y \rangle : \llbracket x \vee y \in I \rrbracket \text{ appears in a sentence in } \Sigma_0\}$. Then Q is finite, and by definition of Σ we have $u = \bigvee_{(x,y) \in Q} (x \wedge y) \in I$. But then

$$u = \bigwedge_{C \subseteq Q} \left(\bigvee_{\langle x,y \rangle \in C} x \vee \bigvee_{\langle x,y \rangle \in Q-C} y \right)$$

Since I is proper, $u \notin I^*$, and hence there exists a subset $C \subseteq Q$ such that

$$\bigvee_{\langle x, y \rangle \in C} x \vee \bigvee_{\langle x, y \rangle \in Q - C} y \notin I^*$$

Fix such a subset C . Define a valuation $v_0 : P \rightarrow \{\top, \perp\}$ by declaring

$$v_0(\llbracket x \vee y \in I \rrbracket) = \begin{cases} \top & \text{if } \langle x, y \rangle \in C \\ \perp & \text{otherwise} \end{cases}$$

and let \widehat{v}_0 be the (unique) extension of the valuation v_0 to $\mathcal{L}(P)$.

Now $\bigvee F_1 \notin I^*$ and $\bigvee F_2 \notin I^*$, where

$$F_1 = \{x \in \mathbb{L} : \langle x, y \rangle \in C \text{ for some } y \in \mathbb{L}\}$$

$$F_2 = \{y \in \mathbb{L} : \langle x, y \rangle \in Q - C \text{ for some } x \in \mathbb{L}\}$$

Since $\bigvee F \in I^*$, we must have $F \not\subseteq F_1$ and $F \not\subseteq F_2$, and so in particular there exist $x_0, y_0 \in F$ such that $\langle x_0, f(x_0) \rangle \notin C$ and $\langle f(y_0), y_0 \rangle \notin Q - C$.

If $\xi(I, F, f) \in \Sigma_0$ then $\langle x_0, f(x_0) \rangle \in Q$. Since $\langle x_0, f(x_0) \rangle \notin C$, we have $v_0(\llbracket x_0 \vee f(x_0) \in I \rrbracket) = \perp$, and so $\widehat{v}_0(\xi(I, F, f)) = \top$. Likewise if $\eta(I, F, f) \in \Sigma_0$ then $\langle f(y_0), y_0 \rangle \in Q$, and since $\langle f(y_0), y_0 \rangle \notin Q - C$ we have $\widehat{v}_0(\llbracket f(y_0) \vee y_0 \in I \rrbracket) = \top$, and so $\widehat{v}_0(\eta(I, F, f)) = \top$.

Thus $\widehat{v}_0^\rightarrow(\Sigma_0) = \{\top\}$, so v_0 is a model of Σ_0 . By the above remark, we may now apply the sentential compactness theorem to deduce that Σ has a model. \square

Denote the model of Σ by $v : P \rightarrow \{\top, \perp\}$. For a proper ideal $I \subseteq \mathbb{L}$ define

$$V_1(I) = \{x \in \mathbb{L} : v(\llbracket x \vee y \in I \rrbracket) = \top \text{ for some } y \in \mathbb{L}\}$$

$$V_2(I) = \{y \in \mathbb{L} : v(\llbracket x \vee y \in I \rrbracket) = \top \text{ for some } x \in \mathbb{L}\}$$

And for $i = 1, 2$ let $V_i^+(I)$ be the ideal generated by $I \cup V_i(I)$.

Lemma 2.11 $V_1^+(I)$ and $V_2^+(I)$ are proper ideals.

Proof Let $F \subseteq V_1(I)$ be a finite subset. For each $x \in F$, there exists $y \in \mathbb{L}$ such that $v(\llbracket x \vee y \in I \rrbracket) = \top$, so let $f(x)$ be some such y . (The axiom of choice is not required here since F is finite.) Then

$$\widehat{v} \left(\bigvee_{x \in F} \neg \llbracket x \vee f(x) \in I \rrbracket \right) = \perp$$

and hence $\bigvee F \notin I^*$, since if $\bigvee F \in I^*$ then $\bigvee_{x \in F} \neg \llbracket x \vee f(x) \in I \rrbracket \in \Sigma$. Since no finite subset of $V_1(I)$ has a join in I^* , $V_1^+(I)$ is proper.

The proof that $V_2^+(I)$ is proper is almost identical. \square

Proposition 2.12 *The sentential compactness theorem implies the almost maximal ideal theorem.*

Proof Define a nondecreasing transfinite sequence of proper ideals of \mathbb{L} as follows:

- $I_0 = \{0\};$
- $I_{\alpha+1} = \begin{cases} V_1^+(I_\alpha) & \text{if } I_\alpha \subsetneq V_1^+(I_\alpha) \\ V_2^+(I_\alpha) & \text{if } I_\alpha = V_1^+(I_\alpha) \subsetneq V_2^+(I_\alpha); \\ \widehat{I} & \text{otherwise} \end{cases}$
- $I_\lambda = \bigcup_{\alpha < \lambda} I_\alpha$ if λ is a limit ordinal.

This sequence is well-defined since the partially ordered set of proper ideals is chain-complete, so that unions of proper ideals are proper, and by Lemma 2.11 the ideals indexed by successor ordinals are proper.

The sequence $\langle I_\alpha : \alpha \in \mathbf{Ord} \rangle$ must eventually be constant with value A say, where A is a proper ideal. Then $A = \widehat{A}$, so it suffices to prove that A is prime so that A will be an almost maximal ideal.

Let $x, y \in \mathbb{L}$ with $x \wedge y \in A$. Then $\llbracket x \vee y \in A \rrbracket \in P$. If $v(\llbracket x \vee y \in A \rrbracket) = \top$ then

$$x \in V_1(A) \subseteq V_1^+(A) = A$$

so $x \in A$. If $v(\llbracket x \vee y \in A \rrbracket) = \perp$ then

$$y \in V_2(A) \subseteq V_2^+(A) = A$$

so $y \in A$. Hence A is almost maximal. □

At long last we are able to prove the final implication, from which Theorem 2.4 will follow.

Proposition 2.13 *The almost maximal ideal theorem implies the Boolean prime ideal theorem.*

Proof Let \mathbb{B} be a Boolean algebra. Then \mathbb{B} is in particular a distributive lattice, so has an almost maximal ideal by **AMIT**. But an almost maximal ideal is prime by definition, so \mathbb{B} has a prime ideal. □

3 Set theory with atoms

Proposition 3.1 *There are no nontrivial \in -automorphisms of the $\text{ZF}(\mathbb{C})$ universe \mathbb{V} .*

Proof Let $f : \mathbb{V} \rightarrow \mathbb{V}$ be an automorphism. If x is a set and $f(y) = y$ for all $y \in x$ then $f(x) = \{f(y) : y \in x\} = \{y : y \in x\} = x$. Therefore f is trivial by \in -induction. \square

The key idea in set theory with atoms is to construct a universe which *does* admit nontrivial \in -automorphisms. This is achieved by allowing for *atoms*, also known in the literature as *urelemente*. Atoms are objects with no set-theoretic structure; they are distinct from sets and are empty in the sense that they contain no elements, but they may be elements of sets. The resulting theory, **ZFA**, is obtained from the axioms of **ZF** by making the following modifications:

- Introduce a unary predicate symbol A , where $A(x)$ will denote ‘ x is an atom’.
- Introduce an *axiom of atoms*, which states that all atoms are empty:

$$\forall x(A(x) \rightarrow \forall y y \notin x)$$

- Modify the empty set axiom to state that there is a *set* which is empty:

$$\exists x(\neg A(x) \wedge \forall y y \notin x)$$

- Modify the axiom of extensionality so that it refers only to sets:

$$\forall x \forall y (\neg A(x) \wedge \neg A(y) \rightarrow (\forall u (u \in x \leftrightarrow u \in y) \rightarrow x = y))$$

This ensures that the atoms do not all collapse onto the empty set.

We will often abuse notation by writing A to denote the class $\{x : A(x)\}$ and $x \in A$ to denote $A(x)$. For the purposes of this essay we will make a further modification:

- Introduce an *atom-set axiom*, which states that there is a set of atoms:

$$\exists x \forall y (y \in x \leftrightarrow A(y))$$

If we did not do this then, without further tampering with the axioms, problems would arise. To illustrate this, consider the powerset axiom:

$$\forall x \exists y \forall z (\forall w (w \in z \rightarrow w \in x) \rightarrow z \in y)$$

For any set x this produces some $y \supseteq A \cup \mathcal{P}(x)$, since all atoms z satisfy $\forall w(w \in z \rightarrow w \in x)$. So if A were a proper class then this axiom would produce a *set* y containing A , and the axiom of separation would give that $A = \{z \in y : A(z)\}$ was a set... a contradiction! Instead of declaring that A be a set we could modify the powerset axiom to state

$$\forall x \exists y \forall z (\neg A(z) \wedge \forall w (w \in z \rightarrow w \in x) \rightarrow z \in y)$$

However, we shall not at any point require that A be a proper class and requiring that it be a set means that we can leave the remaining axioms unchanged.

We can make sense of the notion of rank in **ZFA** in much the same way as we can in **ZF**. For a set (or atom) S , define a cumulative hierarchy $\{\mathcal{P}^\alpha(S) : \alpha \in \mathbf{Ord}\}$ by $\mathcal{P}^0(S) = S$ and

$$\mathcal{P}^{\alpha+1}(S) = \mathcal{P}(\mathcal{P}^\alpha(S)), \quad \mathcal{P}^\lambda(S) = \bigcup_{\alpha < \lambda} \mathcal{P}^\alpha(S), \quad \mathcal{P}^\infty(S) = \bigcup_{\alpha \in \mathbf{Ord}} \mathcal{P}^\alpha(S)$$

where α, λ denote ordinals and λ is a nonzero limit ordinal.

Let $V_\alpha = \mathcal{P}^\alpha(A)$, $\mathbb{V} = \mathcal{P}^\infty(A)$, $K_\alpha = \mathcal{P}^\alpha(\emptyset)$ and $\mathbb{K} = \mathcal{P}^\infty(\emptyset)$.

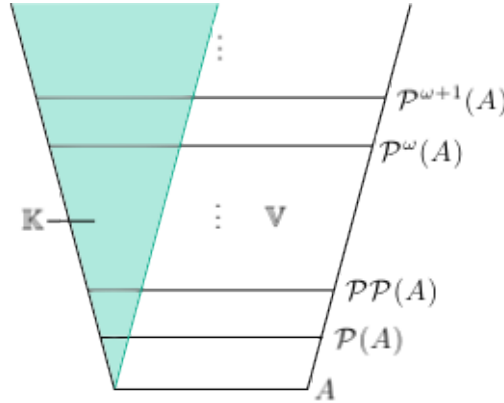


Figure 1: The cumulative hierarchy with atoms

Proposition 3.2 \mathbb{V} is the **ZFA** universe.

Proof Suppose there is some set x with $x \notin \mathbb{V}$. Let $y \in \text{trcl}(x)$ be \in -least such that $y \notin \mathbb{V}$. Then for all $z \in y$ we must have $z \in \mathbb{V}$, and hence there is a least ordinal α_z with $z \in V_{\alpha_z+1}$. But then if $\alpha = \bigcup \{\alpha_z + 1 : z \in y\}$ we must have $z \in V_\alpha$ for each $z \in y$, and so $y \in V_{\alpha+1}$, and hence $y \in \mathbb{V}$, contradicting our assumption. \square

The *rank* of a set $x \in \mathbb{V}$ is the least ordinal α for which $x \in V_{\alpha+1}$. By Proposition 3.2, every set has a rank. Furthermore, \mathbb{K} is a model of **ZF**,

called the *kernel* of \mathbb{V} , and the sets in \mathbb{K} are called *pure sets*. A picture of the ZFA universe is shown in Figure 1.

More generally, if \mathbb{M} is a model of ZFA then $\mathbb{M} \cap \mathbb{K}$ is the kernel of \mathbb{M} . Note also that $\mathbb{K} = H(\neg A)$, or in other words, the pure sets are precisely those which are hereditarily not atoms.

3.1 Fraenkel–Mostowski permutation models

The set of permutations $A \rightarrow A$ forms a group $\text{Aut}(A)$ under composition. Every $f \in \text{Aut}(A)$ extends to a unique \in -automorphism of \mathbb{V} defined recursively on sets by $f(x) = f^\rightarrow(x) = \{f(y) : y \in x\}$. Provided $|A| > 1$, nontrivial \in -automorphisms of \mathbb{V} *will* exist, in stark contrast to Proposition 3.1.

We will exploit this new symmetry to pass to submodels of \mathbb{V} consisting of those sets which are suitably symmetric. By controlling what we mean by ‘suitably symmetric’ we will be able to prove independence results with relatively little effort. For instance, suppose ψ is a formula and ϕ is a formula of the form $\forall x \exists y \phi_0(x, y)$. We will obtain $\text{ZFA} \not\models \psi \rightarrow \phi$ by passing to a submodel in which ψ holds but in which there is a suitably symmetric set x but no suitably symmetric y satisfying $\phi_0(x, y)$.

For a set x and a subgroup $G \leq \text{Aut}(A)$, the (*setwise*) *stabilizer* of x by G is

$$\text{stab}_G(x) = \{f \in \text{Aut}(A) : f(x) = x\}$$

and the *fixator* of x by G , sometimes called its *elementwise stabilizer*, is

$$\text{fix}_G(x) = \{f \in \text{Aut}(A) : f(y) = y \text{ for all } y \in x\} = \bigcap_{y \in x} \text{stab}_G(y)$$

Note that $\text{fix}_G(x) \trianglelefteq \text{stab}_G(x) \leq G$.

A *normal filter* \mathcal{F} (of subgroups of G) is a filter in $\text{Sub}(G)$ which is closed under conjugations, i.e. if $H \in \mathcal{F}$ and $f \in G$ then $f^{-1}Hf \in \mathcal{F}$.

We say a set x is (\mathcal{F} -) *symmetric* if $\text{stab}_G(x) \in \mathcal{F}$. The class of hereditarily \mathcal{F} -symmetric sets is denoted by $\mathbb{M}(A, G, \mathcal{F})$ and called a (*Fraenkel–Mostowski*) *permutation model* of set theory.

Theorem 3.3 $\mathbb{M} = \mathbb{M}(A, G, \mathcal{F})$ is a transitive model of ZFA, $\mathbb{K} \subseteq \mathbb{M}$ and $A \in \mathbb{M}$.

Proof Since \mathbb{M} is defined hereditarily it is a transitive class. Also, G acts trivially on \mathbb{K} , as can be proved by \in -induction: if $f(y) = y$ for all $y \in x$

then

$$f(x) = \{f(y) : y \in x\} = \{y : y \in x\} = x$$

so $\mathbb{K} \subseteq \mathbb{M}$. And $\text{stab}_G(A) = G$ since G is a group of permutations of A .

It remains to prove that \mathbb{M} is a model of ZFA. A proof of this fact, involving the notions of *almost universality* and *Gödel operations*, can be found in [10, Ch. 4, §2]. \square

3.2 Parameter modification

Although a permutation model \mathbb{M} is determined by some parameters A , G and \mathcal{F} , these parameters are not determined by the model itself. Later on it we will need to impose certain conditions on these parameters without any change to the model—in this subsection we prove that this is possible.

Asymmetric atoms

For a given normal filter \mathcal{F} , it may be the case that some atoms are not \mathcal{F} -symmetric. Since the permutation model only sees that which is symmetric, if any atoms are asymmetric then they will be forgotten when passing from \mathbb{V} to \mathbb{M} and it will be as if they never existed at all in the universe. This is made more precise in Proposition 3.4 below.

Proposition 3.4 *Let $\mathbb{M} = \mathbb{M}(A, G, \mathcal{F})$ be a permutation model. Then there exist a subset $A' \subseteq A$ of atoms, a group G' of automorphisms of A' and a normal filter \mathcal{F}' of subgroups of G' , such that*

- (i) $\mathbb{M} \subseteq \mathbb{V}(A') = \mathcal{P}^\infty(A')$;
- (ii) $\mathbb{M} = \mathbb{M}(A', G', \mathcal{F}')$;
- (iii) Each $a \in A'$ is \mathcal{F}' -symmetric;
- (iv) Each $H \in \mathcal{F}'$ is the stabilizer of some $x \in \mathbb{M}$.

Proof Let A' be the collection of all \mathcal{F} -symmetric atoms. If $a \in A - A'$ then $\text{stab}_G(a) \notin \mathcal{F}$ and so $a \notin \mathbb{M}$; so in particular if $x \in \mathbb{V} - \mathbb{V}(A')$ then $x \notin \mathbb{M}$. Hence $\mathbb{M} \subseteq \mathbb{V}(A')$. This proves (i). Define

$$N = \{f \in G : f(a) = a \text{ for all } a \in A'\} \trianglelefteq G$$

Then $G' = G/N$ acts as a group of permutations on A' , and

$$\mathcal{F}/N = \{H/(N \cap H) : H \in \mathcal{F}\}$$

is a normal filter of subgroups of G' and $\mathbb{M}(A', G/N, \mathcal{F}/N) = \mathbb{M}(A, G, \mathcal{F})$.

For $H \leq G$ write $H' = H/N \cap H$, and define

$$\mathcal{F}' = \{H' \in \mathcal{F}/N : H' = \text{stab}_{G'}(x) \text{ for some } x \in \mathbb{M}\}$$

Then \mathcal{F}' is a normal filter:

Upwards closure. Suppose $H' \leq K' \leq G'$ and $H' = \text{stab}_{G'}(x) \in \mathcal{F}'$ for some $x \in \mathbb{M}$. Then $f(x)$ lies in \mathbb{M} for each $f \in G'$, because if $y \in x$ then

$$\text{stab}_{G'}(f(y)) = f \text{stab}_{G'}(y) f^{-1} \in \mathcal{F}'$$

In particular, $z = K'(x) \in \mathbb{M}$. Clearly $K' \subseteq \text{stab}_{G'}(z)$. Conversely, if $f \in \text{stab}_{G'}(z)$ then $f k_1(x) = k_2(x)$ for some $k_1, k_2 \in K'$. Hence $f k_1 k_2^{-1} \in H'$ and so $f \in H' k_2 k_1^{-1} \subseteq K'$. So $K' = \text{stab}_{G'}(z)$.

Closure under intersections. If $H' = \text{stab}_{G'}(x)$ and $K' = \text{stab}_{G'}(y)$ lie in \mathcal{F}' then $H' \cap K' = \text{stab}_{G'}(\langle x, y \rangle) \in \mathcal{F}'$.

Closure under conjugation. If $H' = \text{stab}_{G'}(x) \in \mathcal{F}'$ then

$$f H' f^{-1} = \text{stab}_{G'}(f(x)) \in \mathcal{F}'$$

Now since $\mathcal{F}' \subseteq \mathcal{F}/N$ we have $\mathbb{M}(A', G', \mathcal{F}') \subseteq \mathbb{M}(A', G/N, \mathcal{F}/N) = \mathbb{M}$. The reverse inclusion is immediate from how we defined \mathcal{F}' , so we have equality. Thus (ii) is proved, and (iii) and (iv) follow immediately. \square

The situation of having asymmetric atoms is avoided in some books (e.g. [7], [10]) by encoding into the definition of a normal filter that stabilizers of atoms lie in the filter. However, to avoid any dependence on A in our definition of a normal filter, we will overcome this anomaly by using Proposition 3.4.

Shrinking

Given a group G , a normal filter \mathcal{F} on G and a subgroup $H \in \mathcal{F}$, define the *shrink* \mathcal{F}_H of \mathcal{F} to H by

$$\mathcal{F}_H = \{K \in \mathcal{F} : K \leq H\} = \mathcal{F} \cap \text{Sub}(H)$$

We can shrink the filter giving rise to it without changing the model. Intuitively this is because if a group stabilizes some set then so do all of its subgroups.

Proposition 3.5 *Let A be a set of atoms, G a group of permutations of A and \mathcal{F} a normal filter on G . Then $\mathbb{M}(A, G, \mathcal{F}) = \mathbb{M}(A, H, \mathcal{F}_H)$ for any $H \in \mathcal{F}$.*

Proof If $\text{stab}_H(x) \in \mathcal{F}_H$ then $\text{stab}_H(x) \in \mathcal{F}$, and hence $\text{stab}_G(x) \in \mathcal{F}$ since $\text{stab}_H(x) \leq \text{stab}_G(x)$ and \mathcal{F} is upwards-closed. Conversely, if $\text{stab}_G(x) \in \mathcal{F}$ then $\text{stab}_H(x) = \text{stab}_G(x) \cap H \in \mathcal{F}$ since $H \in \mathcal{F}$ by assumption; but then $\text{stab}_H(x) \in \mathcal{F}_H$ since $\text{stab}_H(x) \leq H$.

Since $\text{stab}_G(x) \in \mathcal{F}$ if and only if $\text{stab}_H(x) \in \mathcal{F}_H$, it follows immediately that $x \in \mathbb{M}(A, G, \mathcal{F})$ if and only if $x \in \mathbb{M}(A, H, \mathcal{F}_H)$. \square

Quotienting out triviality

The final modification we will make later on regards the filter. If the intersection of all the subgroups in \mathcal{F} acts trivially on the sets in \mathcal{M} then we may quotient out by this triviality without changing the model.

Proposition 3.6 *Let $N = \bigcap \mathcal{F}$. Then $N \trianglelefteq G$, and $\mathbb{M}(A, G/N, \mathcal{F}/N) = \mathbb{M}(A, G, \mathcal{F})$.*

Proof Let $g \in G$ and $n \in N$. Then $n \in H$ for all $H \in \mathcal{F}$, so $gng^{-1} \in gHg^{-1}$ for all $H \in \mathcal{F}$. By normality, $gng^{-1} \in H$ for all $H \in \mathcal{F}$, so $gng^{-1} \in N$. Hence $N \trianglelefteq G$. In fact, $N \trianglelefteq H$ for all $H \in \mathcal{F}$, so in particular $N \trianglelefteq \text{stab}_G(x)$ for all $x \in \mathbb{M}$. Since N acts trivially on x , we have

$$\text{stab}_{G/N}(x) = \text{stab}_G(x)/N$$

so $x \in \mathbb{M}(A, G, \mathcal{F})$ if and only if $x \in \mathbb{M}(A, G/N, \mathcal{F}/N)$. \square

3.3 Normal filter–normal ideal duality

In constructing permutation models it will prove useful to be able to characterise the normal filter \mathcal{F} in terms of subsets of A , whenever this is possible. This provides a more convenient setting for working with the wellordering principle, which is equivalent to the axiom of choice by Lemma 1.2. To this end, we say that a nonempty collection \mathcal{I} of subsets of A is a *normal ideal* if it is an ideal of subsets of A which is closed under the action of G , i.e. if $S \in \mathcal{I}$ and $f \in G$ then $f(S) \in \mathcal{I}$. In this way, normal filters give rise to normal ideals and vice versa, and these operations are almost but not quite mutually inverse, in a sense explained below.

If \mathcal{I} is a normal ideal of subsets of A then define

$$\mathcal{I}^* = \{H \leq G : \text{fix}_G(S) \leq H \text{ for some } S \in \mathcal{I}\}$$

This is a normal filter of subgroups of G . Conversely, if \mathcal{F} is a filter, G is a group of permutations of A and $H \leq G$ then define the *H-fixed subset*

$$A^H = \{a \in A : f(a) = a \text{ for all } f \in H\} \subseteq A$$

If \mathcal{F} is a normal filter of subgroups of G then define

$$\mathcal{F}^* = \{S \subseteq A : S \subseteq A^H \text{ for some } H \in \mathcal{F}\}$$

This is a normal ideal of subsets of A .

In passing we note the following result. A *Galois connection* between partially ordered sets \mathbb{P} and \mathbb{P}' is a pair $\langle f, g \rangle$ of monotone maps $f : \mathbb{P} \rightarrow \mathbb{P}'$ and $g : \mathbb{P}' \rightarrow \mathbb{P}$ such that, for all $x \in \mathbb{P}$ and $y \in \mathbb{P}'$, $f(x) \leq y$ if and only if $x \leq g(y)$.

Proposition 3.7 *The assignments $S \mapsto \text{fix}_G(S)$ and $H \mapsto A^H$ define a Galois connection between $\mathcal{P}(A)$ and $\text{Sub}(G)$.*

Proof Let $H \leq G$ and $S \subseteq A$. Then $S \subseteq A^H$ if and only if every element of H fixes each element of S , which occurs if and only if $\text{fix}_G(S) \leq H$. Furthermore, if $K \leq H$ and $T \subseteq S$ then $A^K \supseteq A^H$ and $\text{fix}_G(T) \geq \text{fix}_G(S)$, so the assignments are monotonic. \square

If \mathcal{I} is a normal ideal on A then $x \in \mathbb{M}(A, G, \mathcal{I}^*)$ if and only if there is some $S \in \mathcal{I}$ with $\text{fix}_G(S) \leq \text{stab}_G(x)$. We call such a set S a *support* for x .

Lemma 3.8 *A set $x \in \mathbb{M} = \mathbb{M}(A, G, \mathcal{F})$ has a wellordering in \mathbb{M} if and only if $\text{fix}_G(x) \in \mathcal{F}$.*

Proof The existence of a wellorder of x in \mathbb{M} is equivalent to the existence of an ordinal α and an injection $h : x \rightarrow \alpha$. For each $y \in x$, $h(y) \in \alpha$ lies in the kernel, so is fixed by G . Thus for all $f \in G$ and $y \in x$ we have $f(\langle y, h(y) \rangle) = \langle f(y), h(y) \rangle$. Since h is injective, we thus have $h(f) = f$ if and only if $f(y) = y$ for each $y \in x$. Therefore $f \in \text{stab}_G(h)$ if and only if $f \in \text{fix}_G(x)$, and so $\text{stab}_G(h) = \text{fix}_G(x)$. \square

This fact will be easy to exploit in order to force that certain sets have no wellordering in a given permutation model.

Fraenkel's basic model

Fraenkel's basic model is a very simple permutation model that illustrates how useful the duality between normal filters and normal ideals is. Let A be a countably infinite set of atoms, G be the group of all permutations of A and \mathcal{I} be the set of all finite subsets of A . It is clear that \mathcal{I} is a normal ideal, so we obtain a model $\mathbb{M} = \mathbb{M}(A, G, \mathcal{I}^*)$ of ZFA in the usual way.

Theorem 3.9 $\mathbb{M} \models \neg AC$

Proof Note first that $A \in \mathbb{M}$, since $\text{stab}_G(A) = G \in \mathcal{I}^*$ and each atom is symmetric by definition of \mathcal{I}^* . Now if $S \in \mathcal{I}$ then $A^{\text{fix}_G(S)} = S$ since if $x, y \in A - S$ then the transposition of x and y lies in $\text{fix}_G(S)$. So if $S \in \mathcal{I}$ with $\text{fix}_G(S) \leq \text{fix}_G(A)$ then

$$A = A^{\text{fix}_G(A)} \subseteq A^{\text{fix}_G(S)} = S$$

This contradicts finiteness of S , so $\text{fix}_G(A) \notin \mathcal{I}^*$, and A admits no wellordering in \mathbb{M} by Lemma 3.8. \square

Corollary 3.10 *The axiom of choice is independent from the axioms of ZFA.* \square

Before moving on to a more interesting example of a permutation model, we will explore the duality between normal filters and normal ideals further. In Fraenkel's basic model, it was the fact that the normal filter arose from an ideal that allowed us to break the axiom of choice so easily, but in general normal filters may not arise in this way.

Proposition 3.11

- (i) *Given a normal filter \mathcal{F} on a group G , $\mathcal{F} \supseteq \mathcal{F}^{**}$ and $\mathcal{F}^{***} = \mathcal{F}^*$.*
- (ii) *Given a normal ideal \mathcal{I} on a set A , $\mathcal{I} \subseteq \mathcal{I}^{**}$ and $\mathcal{I}^{***} = \mathcal{I}^*$.*

Proof

- (i) Let $H \in \mathcal{F}^{**}$. Then there exists $S \in \mathcal{F}^*$ with $\text{fix}_G(S) \leq H$, and for this S there is some $K \in \mathcal{F}$ with $S \subseteq A^K$. But then $K \leq \text{fix}_G(S) \leq H$, and so $H \in \mathcal{F}$ by upwards-closure.

The first part of (ii) applied to \mathcal{F}^* gives $\mathcal{F}^* \subseteq \mathcal{F}^{***}$. If $S \in \mathcal{F}^{***}$ then $\text{fix}_G(S) \leq H$ for some $H \in \mathcal{F}^{**}$; but then $H \in \mathcal{F}$ since $\mathcal{F}^{**} \subseteq \mathcal{F}$, so $S \in \mathcal{F}^*$ and $\mathcal{F}^* = \mathcal{F}^{***}$.

- (ii) Let $S \in \mathcal{I}$, then $\text{fix}_G(S) \in \mathcal{I}^*$, so $A^{\text{fix}_G(S)} \in \mathcal{I}^{**}$. But $S \subseteq A^{\text{fix}_G(S)}$, so $S \in \mathcal{I}^{**}$ by downwards-closure.

The first part of (i) applied to \mathcal{I}^* gives $\mathcal{I}^{***} \subseteq \mathcal{I}^*$. If $H \in \mathcal{I}^*$ then $\text{fix}_G(S) \leq H$ for some $S \in \mathcal{I}$; but then $S \in \mathcal{I}^{**}$ since $\mathcal{I} \subseteq \mathcal{I}^{**}$, so $H \in \mathcal{I}^{***}$ and $\mathcal{I}^* = \mathcal{I}^{***}$. \square

A normal filter \mathcal{F} (resp. normal ideal \mathcal{I}) is *reflexive* if $\mathcal{F} = \mathcal{F}^{**}$ (resp. $\mathcal{I} = \mathcal{I}^{**}$). Note that if a normal filter (resp. ideal) \mathcal{F} is irreflexive then there does not exist a normal ideal (resp filter) \mathcal{I} such that $\mathcal{F} = \mathcal{I}^*$. The following two examples demonstrate that Proposition 3.11 is the best we can hope for, in that irreflexive normal filters and ideals exist.

Example Let G be the group of all permutation of \mathbb{Z} acting on \mathbb{R} by $f(n + \delta) = f(n) + \delta$ for all $f \in G$, $n \in \mathbb{Z}$ and $0 \leq \delta < 1$; that is, G acts on \mathbb{R} by permuting intervals of the form $[n, n + 1)$ for $n \in \mathbb{Z}$. Let \mathcal{I} be the normal ideal of all finite subsets of \mathbb{R} . Then $[0, 1) \in \mathcal{I}^{**}$ since $[0, 1) = \mathbb{R}^{\text{fix}_G(\{0\})}$, but $[0, 1)$ is infinite, so $[0, 1) \notin \mathcal{I}$. Hence $\mathcal{I} \neq \mathcal{I}^{**}$.

Example For $X \subseteq \mathbb{Z}$ let S_X denote the group of permutations of X which fix all but finitely many points and A_X denote the subgroup of S_X consisting of all even permutations.

$$\mathcal{F} = \{S_{\mathbb{Z}-E} : E \subseteq \mathbb{Z} \text{ finite}\} \cup \{A_{\mathbb{Z}-E} : E \subseteq \mathbb{Z} \text{ finite}\}$$

Write $G = S_{\mathbb{Z}}$, and note that $S_{\mathbb{Z}-E} = \text{fix}_G(E)$ for all $E \subseteq \mathbb{Z}$. Now \mathcal{F} is closed under intersections, since

$$S_{\mathbb{Z}-E} \cap S_{\mathbb{Z}-F} = S_{\mathbb{Z}-(E \cup F)}$$

$$S_{\mathbb{Z}-E} \cap A_{\mathbb{Z}-F} = A_{\mathbb{Z}-(E \cup F)}$$

$$A_{\mathbb{Z}-E} \cap A_{\mathbb{Z}-F} = A_{\mathbb{Z}-(E \cup F)}$$

and it is closed upwards by elementary results in group theory. Closure under conjugation follows from the fact that $fS_{\mathbb{Z}-E}f^{-1} = S_{\mathbb{Z}-f(E)}$ and $fA_{\mathbb{Z}-E}f^{-1} = A_{\mathbb{Z}-f(E)}$. Hence \mathcal{F} is a normal filter. However $A_{\mathbb{Z}} \notin \mathcal{F}^{**}$ since $\text{fix}_G(\mathbb{Z}^{\mathbb{A}_{\mathbb{Z}}}) = G$ and $\text{fix}_G(\mathbb{Z}^{\text{fix}_G(\{n\})}) = \text{fix}_G(\{n\})$, both of which contain odd permutations. Hence $\mathcal{F} \neq \mathcal{F}^{**}$.

4 Mostowski's ordered model

By introducing an order structure on the set of atoms, Mostowski's ordered model allows us to refine the result obtained from Fraenkel's model to prove not just that the axiom of choice is independent from the axioms of ZFA (Lemma 4.1), but in fact that the axiom of choice is independent from the ordering principle in the presence of the ZFA axioms (Theorem 4.4). With a bit more work, we will see that the Boolean prime ideal theorem also holds in Mostowski's ordered model (Theorem 4.7), therefore proving that the axiom of choice is independent from the Boolean prime ideal theorem in the presence of the ZFA axioms.

Let A denote the set of atoms, which is assumed to be countably infinite, and let $<_A$ be an ordering on A such that $\langle A, <_A \rangle$ is order-isomorphic to $\langle \mathbb{Q}, < \rangle$. Let G be the group of all $<_A$ -preserving permutations of A and let \mathcal{I} be the normal ideal of all finite subsets of A . Again, we get an induced permutation model $\mathbb{M} = \mathbb{M}(A, G, \mathcal{I}^*)$.

Lemma 4.1 $\mathbb{M} \models A$ has no wellordering.

Proof Let S be a finite subset of A . Let $h : \mathbb{Q} \rightarrow A$ be an order-isomorphism and let $\tilde{S} = h^{\leftarrow}(S) \subseteq \mathbb{Q}$. Suppose that $\max(\tilde{S}) \leq 0$; otherwise replace f by $f - \max(\tilde{S})$. Define $\tilde{f} : \mathbb{Q} \rightarrow \mathbb{Q}$ by

$$\tilde{f}(x) = \begin{cases} x & \text{if } x \leq 0 \\ 2x & \text{if } x \geq 0 \end{cases}$$

Then $f = h \circ \tilde{f} \circ h^{-1}$ is a nontrivial $<_A$ -preserving permutation of A which fixes each $s \in S$. \square

So the axiom of choice fails in \mathbb{M} . In order to prove the independence of the ordering principle from the axiom of choice, we need to find a way of linearly ordering each set in \mathbb{M} . To do so, some preliminary results are needed.

Lemma 4.2 Each $x \in \mathbb{M}$ has a least support S_x , and the class of all pairs $\langle x, S_x \rangle$ for $x \in \mathbb{M}$ is \mathcal{I}^* -symmetric.

Proof Let S and T be supports of x , and let $U = S \cap T$.

Clearly $\text{fix}_G(S), \text{fix}_G(T) \leq \text{fix}_G(U)$. Moreover, if $f \in \text{fix}_G(U)$ then we can find $g_1, \dots, g_n \in \text{fix}_G(S) \cup \text{fix}_G(T)$ with $g_n \circ \dots \circ g_1 = f$. Intuitively, this is done as follows. Partition $A - U$ into $|U| + 1$ sets in the obvious way.

Taking each partition in turn, alternately slide the elements of S and T in the partition along (by elements of $\text{fix}_G(T)$ and $\text{fix}_G(S)$, respectively) until they and all the points between them lie at their images under f . This can be done in finitely many steps since $S \cup T$ is finite. It follows that $\text{fix}_G(U)$ is equal to the group generated by $\text{fix}_G(S)$ and $\text{fix}_G(T)$ and hence U is a support of x .

This implies that any finite intersection of supports is a support. Now if \mathcal{S} is the collection of all supports of x then there is some finite subcollection $\mathcal{S}' \subseteq \mathcal{S}$ such that $\bigcap \mathcal{S}' = \bigcap \mathcal{S} = S_x$; for instance, pick an arbitrary $S \in \mathcal{S}$ and let $\mathcal{S}' = \mathcal{S} \cap \mathcal{P}(S)$. Hence x has a least support.

To prove the last part, note that if $f \in G$, $S \in \mathcal{I}$ and $x \in \mathbb{M}$ then

$$\text{fix}_G(f(S)) = f \text{fix}_G(S) f^{-1} \quad \text{and} \quad \text{stab}_G(f(x)) = f \text{stab}_G(x) f^{-1}$$

so S_x is the least support of x if and only if $f(S_x)$ is the least support of $f(x)$. \square

Lemma 4.3 *There is a symmetric injective function-class $\iota : \mathbb{M} \rightarrow \mathbf{Ord} \times \mathcal{I}$.*

Proof The orbits $G(x) = \{f(x) : f \in G\}$ are pairwise disjoint and symmetric, so can be enumerated by ordinals, and furthermore this enumeration is symmetric. Let α_x be the ordinal assigned to the orbit $G(x)$, and define $\iota(x) = \langle \alpha_x, S_x \rangle$. Then ι is symmetric, since $f(S_x) = S_{f(x)}$ and $f(x) \in G(x)$ so $\alpha_{f(x)} = \alpha_x$. Furthermore, ι is injective: if $\iota(x) = \iota(y)$ then $S_x = S_y$ and $y \in G(x)$. Let $y = f(x)$ for $f \in G$. Then $S_x = S_{f(x)} = f(S_x)$, and so $f \in S_x$. But that means that $f(x) = x$, so $y = x$. \square

Theorem 4.4 $\mathbb{M} \models \text{Every set has a linear ordering.}$

Proof Note first that $<_A$ is symmetric, since by definition each $f \in G$ respects $<_A$. Furthermore, \mathcal{I} can be ordered lexicographically since it is a set of finite subsets of a linearly ordered set; hence $\mathbf{Ord} \times \mathcal{I}$ can be ordered lexicographically (by $<_{\mathbf{Ord} \times \mathcal{I}}$, say). Hence we can define the relation-class $<_{\mathbb{M}}$ on \mathbb{M} by $x <_{\mathbb{M}} y$ if and only if $\iota(x) <_{\mathbf{Ord} \times \mathcal{I}} \iota(y)$. Then $<_{\mathbb{M}}$ is a linear ordering of \mathbb{M} , and so descends to a linear ordering of each set $x \in \mathbb{M}$ by restriction. \square

Corollary 4.5 *The axiom of choice is independent from the ordering principle in the presence of the axioms of ZFA.* \square

Now we will prove that the Boolean prime ideal theorem holds in \mathbb{M} . It may seem strange to do this because this result is stronger than Theorem 4.4; however, its proof is far more involved and uses different methods.

By Theorem 2.4 it suffices to prove that the consistency principle holds in \mathbb{M} , and by Lemma 1.1 we need only consider messes on sets. Given a binary mess $\mathcal{M} \in \mathbb{M}$ on a set $X \in \mathbb{M}$, define $\widehat{\mathcal{M}}$ and \widehat{X} as follows. Let $S_{\mathcal{M}}$ be the least support of \mathcal{M} and, for each $x \in X$, let

$$\mathcal{O}(x) = \text{fix}_G(S_{\mathcal{M}})(x) = \{f(x) : f \in \text{fix}_G(S_{\mathcal{M}})\}$$

$$\widehat{X} = \mathcal{O}^{\rightarrow}(X) = \{\mathcal{O}(x) : x \in X\}$$

Since $\text{fix}_G(S_{\mathcal{M}}) \leq \text{stab}_G(X)$ we have $\mathcal{O}(x) \subseteq X$ for each $x \in X$, and hence $\widehat{X} \subseteq \mathcal{P}(X)$.

Given a finite subset $F \subseteq \widehat{X}$, a subset $E \subseteq F$, and a finite subset $P \subseteq \bigcup F$, we say P *admits* E if $\langle P, \mathcal{O}^{\leftarrow}(E) \rangle \in \mathcal{M}$. Let $\widehat{\mathcal{M}}$ be the collection of all pairs $\langle F, E \rangle$ with $F \subseteq \widehat{X}$ finite and $E \subseteq F$ such that every subset of $\bigcup F$ admits E .

Lemma 4.6 $\widehat{\mathcal{M}}$ is a mess on \widehat{X} .

Idea of proof A full proof can be found in [7, Theorem 7.16] or [10, Theorem 7.1], with a slight change in language: pairs $\langle F, E \rangle$ correspond with finite partial functions $g : X \rightarrow 2$ with $F = \text{dom}(g)$ and $E = g^{\leftarrow}(1)$.

It suffices to prove that for each finite $F \subseteq \widehat{X}$ there is a subset $E \subseteq F$ with $\langle F, E \rangle \in \widehat{\mathcal{M}}$. The proof splits into the cases where $S_{\mathcal{M}}$ is empty and nonempty and appeals to combinatorial arguments about n -sets, i.e. sets that can be partitioned into a union of n sets of the form $f_i^{\rightarrow}(P_0)$ for $f_i \in G$ and P_0 fixed, and *types* of n -sets, i.e. equivalence classes of n -sets under the relation of differing by an element of G . \square

Theorem 4.7 *The consistency principle holds in \mathbb{M} .*

Proof Let $\mathcal{M} \in \mathbb{M}$ be a mess on a set $X \in \mathbb{M}$. Working outside of \mathbb{M} —in particular, assuming the axiom of choice holds in our metatheory—we can find a subset $\widehat{C} \subseteq \widehat{X}$ which is consistent with $\widehat{\mathcal{M}}$. Now define

$$C = \{x \in X : \mathcal{O}(x) \in \widehat{C}\} \subseteq X$$

Then C is symmetric, so $C \in \mathbb{M}$, and C must be consistent with \mathcal{M} . \square

Corollary 4.8 *The Boolean prime ideal theorem is independent from the axiom of choice in the presence of the axioms of ZFA.*

Proof $\mathbb{M} \models \neg\text{AC}$ by Lemma 4.1 and $\mathbb{M} \models \text{CP}$ by Theorem 4.7. By Theorem 2.4, CP is equivalent to BPIT, and so $\mathbb{M} \models \text{BPIT} + (\neg\text{AC})$. \square

5 Topological groups and their actions

This short section builds up some theory necessary to state and prove the main result of the essay.

5.1 Extremely amenable groups

A G -set is a set X equipped with an action $G \curvearrowright X$. A *topological group* is a topological space G equipped with a group structure whose operations

$$(-) \cdot (-) : G \times G \rightarrow G, \quad (-)^{-1} : G \rightarrow G$$

are continuous with respect to the topology.

A *neighbourhood base* \mathcal{N} of $x \in X$ is a collection of subsets of G which contain x , such that if $x \in U \subseteq G$ then U is open if and only if $N \subseteq U$ for some $N \in \mathcal{N}$. A topological group G has *small open subgroups* if and only if every open neighbourhood of the identity contains an open subgroup, i.e. the open subgroups form a neighbourhood base of 1.

For a topological group G and a normal filter \mathcal{F} of subgroups of G , define

$$\mathcal{F}^\dagger = \{U \subseteq G : \forall g \in U \exists H \in \mathcal{F} \ gH \subseteq U\}$$

Conversely, for a topology \mathcal{O} on G , define

$$\mathcal{O}^\dagger = \{H \leq G : H \in \mathcal{O}\} = \{\mathcal{O}\text{-open subgroups of } G\}$$

Proposition 5.1 \mathcal{F}^\dagger is a topology and \mathcal{O}^\dagger is a normal filter. Moreover $\mathcal{F}^{\dagger\dagger} = \mathcal{F}$, and G has small open subgroups if and only if $\mathcal{O} \subseteq \mathcal{O}^{\dagger\dagger}$.

Proof Trivially $\emptyset \in \mathcal{F}^\dagger$ and $G \in \mathcal{F}^\dagger$. Let $U, V \in \mathcal{F}^\dagger$. If $g \in U \cap V$ then there exist $H, K \leq G$ such that $gH \subseteq U$ and $gK \subseteq V$. Then $gH \cap gK = g(H \cap K) \subseteq U \cap V$, and $H \cap K$ is a subgroup, so $U \cap V \in \mathcal{F}^\dagger$. Let $\{U_i : i \in I\}$ be a collection of subsets of G , all of which we may assume to be nonempty, and let $U = \bigcup_{i \in I} U_i$. If $g \in U$ then $g \in U_j$ for some $j \in I$. Then there exists $H \leq G$ such that $gH \subseteq U_j$; so $gH \subseteq U$, so $I \in \mathcal{F}^\dagger$. Hence \mathcal{F}^\dagger is a topology on G .

\mathcal{O}^\dagger is nonempty since G is an open subgroup. It is closed under pairwise intersections since if H and K are open subgroups then $H \cap K$ is an open subgroup. It is closed upwards since if $K \leq H \leq G$ with K open then $H = \bigcup_{h \in H} hK$ and each coset hK is open by continuity of the action of G on itself, so H is open. Likewise, if $g \in G$ and $H \leq G$ is open then $gHg^{-1} \leq G$ is open by continuity. Hence \mathcal{O}^\dagger is a normal filter.

If $H \in \mathcal{F}$ then $hH \subseteq H$ for all $h \in H$, so $H \in \mathcal{F}^{\dagger\dagger}$. Conversely if $K \in \mathcal{F}^{\dagger\dagger}$ then there is some $H \in \mathcal{F}$ with $1H = H \subseteq K$, so $K \in \mathcal{F}$ by upwards-closure.

Now $\mathcal{O}^{\dagger\dagger}$ is precisely the set of all subsets $U \subseteq G$ such that for all $g \in U$ there is some $H \leq G$ with $gH \subseteq U$. This occurs if and only if $H \subseteq g^{-1}U$. Now U is open if and only if $g^{-1}U$ is open, so $\mathcal{O} \subseteq \mathcal{O}^{\dagger\dagger}$ if and only if G has small open subgroups. \square

Recall from Proposition 3.6 that without changing our model we may replace G by G/N where $N = \bigcap \mathcal{F}$. Together with the following result, this means that we may replace G by a quotient of G and thus assume that G is Hausdorff.

Proposition 5.2 *Let G be a topological group with topology \mathcal{O} and suppose that G has small open subgroups. Then $\bigcap \mathcal{O}^{\dagger} = \{1\}$ if and only if G is Hausdorff.*

Proof If $1 \neq g \in \bigcap \mathcal{O}^{\dagger}$ then $g \in H$ for each $H \in \mathcal{O}^{\dagger}$. In particular, if $U \subseteq G$ is open and contains 1 then $g \in U$, so 1 and g cannot be separated by open sets, so G is not Hausdorff.

Conversely, suppose G is Hausdorff. Then for each $g \in G$ there exists an open subset $U_g \subseteq G$ containing 1 but not g , and hence there exists an open subgroup H_g not containing g . Then $\bigcap \mathcal{O}^{\dagger} \subseteq \bigcap_{g \in G} H_g = \{1\}$. \square

If G is a topological group and X is a topological space then an action $G \curvearrowright X$ is *continuous* if it is continuous as a map $G \times X \rightarrow X$ of topological spaces. A topological space X equipped with a continuous G -action is called a *G -space*.

A compact Hausdorff G -space is called a *G -flow*. A group G is *extremely amenable* if for each G -flow X there is a point $x^* \in X$ for which $g \cdot x^* = x^*$ for all $g \in G$, called a *fixed point* of the G -flow.

Extremely amenable groups are at one end of our correspondence; as such we need to make sure no damage is done when passing to a Hausdorff quotient using Propositions 3.6 and 5.2.

Proposition 5.3 *Let G be a topological group with topology \mathcal{O} and write $N = \bigcap \mathcal{O}^{\dagger}$. Then G has small open subgroups if and only if G/N does. Moreover, if G has small open subgroups, then G is extremely amenable if and only if G/N is extremely amenable.*

Proof The quotient topology on G/N is precisely $\mathcal{O}/N = \{U/N : U \in \mathcal{O}\}$, where we take the topological quotient. Furthermore, $\mathcal{O} \subseteq \mathcal{O}^{\dagger\dagger}$ if and only if $\mathcal{O}/N \subseteq \mathcal{O}^{\dagger\dagger}/N = (\mathcal{O}/N)^{\dagger\dagger}$. By Proposition 5.1, G has small open subgroups if and only if G/N does.

Let X be a G -flow. For each open $V \subseteq X$, the set

$$U(V) = \{g \in G : g^\rightarrow(V) \subseteq V\}$$

is open in G , and hence contains some $H \in \mathcal{O}^\dagger$. In particular, $N \subseteq U(V)$, so for all $n \in N$ and open $V \subseteq X$, $n^\rightarrow(V) \subseteq V$. Since X is Hausdorff, N acts trivially on X .

Thus there is an exact correspondence between actions of G on X and actions of G/N on X given by $(gN) \cdot x = g \cdot x$. In particular, x^* is a fixed point for $G \curvearrowright X$ if and only if it is a fixed point for the corresponding action $G/N \curvearrowright X$. \square

Lemma 5.4 *A G -flow X has a fixed point if and only if for each $n \in \mathbb{N}$, continuous $f : X \rightarrow \mathbb{R}^n$, $\varepsilon > 0$ and finite subset $F \subseteq G$, there exists $x \in X$ with*

$$\|f(x) - f(g \cdot x)\| \leq \varepsilon \text{ for all } g \in F$$

where $\|-\|$ denotes the Euclidean norm on \mathbb{R}^n .

Proof (\Rightarrow) This direction is trivial by setting $x = x^*$.

(\Leftarrow) Given f, ε, F as in the statement of the lemma, put

$$A_{f, \varepsilon, F} = \{x \in X : \|f(x) - f(g \cdot x)\| \leq \varepsilon \text{ for all } g \in F\}$$

Then each $A_{f, \varepsilon, F}$ is closed, and hence compact since X is compact.

Now if $\{A_{f_j, \varepsilon_j, F_j} : 1 \leq j \leq m\}$ is any finite collection of such sets, with $f_j : X \rightarrow \mathbb{R}^{n_j}$ for each j , then putting

$$\bar{f} = (f_1, \dots, f_m) : X \rightarrow \mathbb{R}^{n_1 + \dots + n_m}, \quad \bar{\varepsilon} = \min_{1 \leq j \leq m} \varepsilon_j, \quad \bar{F} = \bigcup_{j=1}^m F_j$$

gives $\emptyset \neq A_{\bar{f}, \bar{\varepsilon}, \bar{F}} \subseteq \bigcap_{j=1}^m A_{f_j, \varepsilon_j, F_j}$; that is, the intersection of any finite collection of the sets $A_{f, \varepsilon, F}$ is nonempty.

By compactness the intersection A of *all* the $A_{f, \varepsilon, F}$ is nonempty. Let $x^* \in A$ and suppose it is not a fixed point. Choose $g \in G$ with $x^* \neq g \cdot x^*$. By Lemma 1.3 there is a continuous map $f : X \rightarrow \mathbb{R}$ with $f(x^*) = 0$ and $f(g \cdot x^*) = 1$; but this contradicts $x^* \in A$.

So any point in A is fixed. In particular, $A \neq \emptyset$, so a fixed point exists. \square

5.2 Ramsey filters

A G -set X has the *Ramsey property* if for every 2-colouring $c : X \rightarrow 2$ and every finite $F \subseteq X$ there exists $g \in G$ such that $g \cdot F$ is monochromatic, i.e. $c(g \cdot x) = c(g \cdot x')$ for all $x, x' \in F$. In this case we call the action of G on X a *Ramsey action* and we call X a *Ramsey G -set*.

A subgroup $H \leq G$ is a *Ramsey subgroup* of G if the set G/H of left-cosets of H in G is a Ramsey G -set under the usual action $g \cdot (xH) = (gx)H$. To spell it out, this means that for every 2-colouring $c : G/H \rightarrow 2$ and every finite set F of cosets of H in G , some G -translation $g \cdot F$ is monochromatic.

We say a normal filter \mathcal{F} of subgroups of a group G is a *Ramsey filter* (as *witnessed* by H) if it contains a subgroup H for which whenever $K \leq H$ with $K \in \mathcal{F}$, K is a Ramsey subgroup of H .

When we come to prove the correspondence between models and groups in Section 6, our strategy will be to go via Ramsey filters. It will become apparent that our definition is quite tight and for our purposes it will be convenient to have a weaker definition. The following proposition says we can do that at no cost.

Lemma 5.5 *Fix $k \geq 2$. A G -set X has the Ramsey property if and only if for every finite subset $F \subseteq X$ there exists a finite subset $Y \subseteq X$ such that, whenever $c : Y \rightarrow k$ is a k -colouring of Y , there exists $g \in G$ with $g \cdot F \subseteq Y$ and $g \cdot F$ is monochromatic.*

Proof Denote the latter statement, for a given k , by $R(k)$.

(\Rightarrow) is clear: suppose that X has the Ramsey property; then since a 2-colouring is a k -colouring, we may simply take $Y = g \cdot F$ and consider the restriction of said colouring to Y .

(\Leftarrow) is proved by induction on k , using a compactness argument in the base case. Take $\llbracket x \text{ gets colour } i \rrbracket$ for $x \in X$ and $i \in 2$ as sentential variables, encoded as elements of $X \times 2$ say, and for each 2-colouring $c : X \rightarrow 2$ define a valuation $v_c : X \times 2 \rightarrow \{\top, \perp\}$ by

$$v_c(\llbracket x \text{ gets colour } i \rrbracket) = \begin{cases} \top & \text{if } c(x) = i \\ \perp & \text{otherwise} \end{cases}$$

Let Σ be the set of sentences of the form

$$\begin{aligned} &\llbracket x \text{ gets colour } 0 \rrbracket \vee \llbracket x \text{ gets colour } 1 \rrbracket \\ &\neg(\llbracket x \text{ gets colour } 0 \rrbracket \wedge \llbracket x \text{ gets colour } 1 \rrbracket) \end{aligned}$$

$$\neg \left(\bigwedge_{x \in F} \llbracket g \cdot x \text{ gets colour } 0 \rrbracket \vee \bigwedge_{x \in F} \llbracket g \cdot x \text{ gets colour } 1 \rrbracket \right)$$

for $x \in X$ and $g \in G$. In turn, these mean that each $x \in X$ has a colour, that no $x \in X$ has two colours, and that no G -translation $g \cdot F$ of F is monochromatic.

Let F be a finite subset that witnesses the failure of $R(2)$. Then any finite subset of Σ has a model v_c for some colouring c , and hence Σ has a model by compactness. But this is precisely the statement that F witnesses the failure of the Ramsey property.

Now suppose $k > 2$ and that $R(k')$ holds for all $2 \leq k' < k$. Fix a finite subset $F \subseteq X$. We may choose Y that satisfies $R(k-1)$ using F ; and without loss of generality $Y \supseteq F$ (if not, replace it by $Y \cup F$). By $R(2)$ with Y in the place of F , there exists a finite subset $Z \supseteq Y$ such that every 2-colouring of Z makes some subset of the form $g \cdot Y$ monochromatic.

Let $c : Z \rightarrow k$ be a k -colouring, and define $c_0 : Z \rightarrow 2$ by

$$c_0(z) = \begin{cases} 0 & \text{if } c(z) \neq k-1 \\ 1 & \text{if } c(z) = k-1 \end{cases}$$

By assumption there is some $g \in G$ such that c_0 is constant on $g \cdot Y$.

If $c_0^\rightarrow(g \cdot Y) = \{1\}$ then $c^\rightarrow(g \cdot Y) = \{k-1\}$, so in particular c is constant on $g \cdot F$ and we're done.

If not then $c_0^\rightarrow(g \cdot Y) = \{0\}$ and $c^\rightarrow(g \cdot Y) \subseteq k-1$, so there is a $(k-1)$ -colouring $c_1 : Y \rightarrow k-1$ defined by $c_1(y) = c_0(g \cdot y)$. By $R(k-1)$ there is some $g' \in G$ such that $g' \cdot F \subseteq Y$ and c_1 is constant on $g' \cdot F$; but then c is constant on $(gg') \cdot F$. \square

From now on, the Ramsey property will interchangeably refer to the Ramsey property as originally defined and the statement $R(k)$ for an arbitrary k as in Lemma 5.5.

6 The correspondence

We are now ready to state the long-promised correspondence. Its proof relies on two theorems (6.2 and 6.5), whose proofs occupy the remainder of the section. Throughout this section, G will be a topological group of permutations of a set A of atoms, \mathcal{F} will be the normal filter of open subgroups of G , and $\mathbb{M} = \mathbb{M}(A, G, \mathcal{F})$ will be the corresponding permutation model of ZFA.

Theorem 6.1 *Suppose G is Hausdorff and has small open subgroups, that all atoms are \mathcal{F} -symmetric and that each $H \in \mathcal{F}$ stabilizes some set in \mathbb{M} . Then $\mathbb{M} \models \text{BPIT} + (\neg\text{AC})$ if and only if G is nontrivial and extremely amenable (with a slight caveat).*

Proof

$$\begin{array}{ccc} G \text{ is nontrivial and} & \xleftrightarrow{(6.2)} & \mathcal{F} \text{ is a Ramsey} \\ \text{extremely amenable} & & \text{filter of subgroups} \\ & & \text{of } G \text{ and } \{1\} \notin \mathcal{F} \end{array} \xleftrightarrow{(6.5)} \mathbb{M} \models \text{BPIT} + (\neg\text{AC})$$

The caveat is made precise in Theorems 6.2 and 6.5 and discussed in Section 6.1. \square

Theorem 6.2 *Suppose G is Hausdorff and has small open subgroups. Then the following are equivalent:*

- (i) G is nontrivial and extremely amenable;
- (ii) \mathcal{F} is a Ramsey filter as witnessed by G , and $\{1\} \notin \mathcal{F}$.

Proof

Suppose (i) holds, so that G is nontrivial and extremely amenable. Let $H \leq G$ be an open subgroup, $c : G/H \rightarrow 2$ be a 2-colouring of G/H , and $A \subseteq G/H$ be a finite subset.

G acts on $Y = 2^{G/H}$ by $(g \cdot p)(xH) = p(g^{-1}xH)$ for $p \in Y$, $g \in G$ and $xH \in G/H$, and this action turns Y into a G -flow. Note also that the colouring c lies in Y . Let $X = \overline{G(c)}$ be the closure in Y of the orbit of c under the action of G . Then X is a G -flow and there is a fixed point $p^* \in X$.

For any $g \in G$ we have $p^*(gH) = (g^{-1} \cdot p^*)(H) = p^*(H)$. But G acts transitively on G/H , so $p^* : G/H \rightarrow 2$ is constant, say $p^*(gH) = i$ for all $gH \in G/H$, where $i = 0$ or 1 .

Since $p^* \in \overline{G(c)}$, there is a $g \in G$ for which $(g \cdot c)(aH) = p^*(aH) = i$ for all $aH \in A$. Hence H is a Ramsey subgroup of G , so (ii) holds.

Now suppose that (ii) holds. Then \mathcal{F} satisfies the corresponding definition of a Ramsey filter (witnessed by G) referring to right-cosets H/G rather than left-cosets G/H , where the action $G \curvearrowright H/G$ is given by $g \cdot Hh = H(hg^{-1})$.

Let X be a G -flow, $f : X \rightarrow \mathbb{R}^n$ be a continuous map, $\varepsilon > 0$ and $F \subseteq G$ be a finite subset. Let $H \subseteq G$ be an open neighbourhood of $1 \in G$ for which

$$\|f(x) - f(h \cdot x)\| \leq \frac{\varepsilon}{3} \quad \text{for all } h \in H \text{ and } x \in X$$

Since G has small open subgroups, we can suppose without loss of generality that H is a subgroup of G .

Since $f^{-1}(X) \subseteq \mathbb{R}^n$ is compact, we can partition it into m sets A_0, \dots, A_{m-1} of diameter $\leq \frac{\varepsilon}{3}$. Fix $x_0 \in X$, and for each $0 \leq i < m$ let

$$K_i = \{g \in G : f(g \cdot x_0) \in A_i\}$$

and

$$H_i = \{hk : h \in H, k \in K_i\} = \bigcup_{k \in K_i} Hk$$

We can identify H_i with the subset $\{Hk : k \in K_i\}$ of H/G . Also, $\bigcup_i H_i = G/H$, and so there is an m -colouring $c : G/H \rightarrow m$ such that $c^{-1}(\{i\}) \subseteq H_i$. But then, since \mathcal{F} is a Ramsey filter, there is $0 \leq i < m$ and $g \in G$ with $(F \cup \{1\})g \subseteq H_i$. Set $x = g \cdot x_0$ for this g .

For $y \in F$, choose $h \in H$ with $hyg \in K_i$. Then $f(hyg \cdot x_0) = f(hy \cdot x) \in A_i$. Then $\|f(hy \cdot x) - f(y \cdot x)\| \leq \frac{\varepsilon}{3}$, and so $f(y \cdot x)$ lies in the $\frac{\varepsilon}{3}$ -neighbourhood of A_i . But $1 \in F$, so $\|f(x) - f(y \cdot x)\| \leq \varepsilon$.

It follows from Lemma 5.4 that X has a fixed point for $G \curvearrowright X$, and so (i) holds. \square

To prove Theorem 6.5 we introduce some notation and prove some preliminary results. The idea is similar to that of Theorem 2.12.

Suppose now that each subgroup in \mathcal{F} stabilizes some set in \mathbb{M} ; this is possible by Proposition 3.4. Fix a subgroup $H \in \mathcal{F}$.

Let $K, L \in \mathcal{F}$ be subgroups of H and consider H/L as a K -set. Suppose F is a finite subset of some K -orbit $T \subseteq H/L$ which is a counterexample to the Ramsey property, in the sense that there is a finite subset $Y \subseteq T$ such that for all $k \in K$, Y can be partitioned into two sets Y_{\top} and Y_{\perp} , neither wholly containing kF .

For such F, K, L define the $\langle F, K, L \rangle$ -orbit $X_{\langle F, K, L \rangle}$ to be some set isomorphic (as an H -set) to H/L which is disjoint from $X_{\langle F', K', L' \rangle}$ for $\langle F, K, L \rangle \neq \langle F', K', L' \rangle$. For a given F , let \bar{F} be the image of F under the isomorphism $H/L \cong X_{\langle F, K, L \rangle}$.

Let

$$X = \bigcup_{\langle F, K, L \rangle} X_{\langle F, K, L \rangle}$$

with H -action induced by the isomorphisms $X_{\langle F, K, L \rangle} \cong H/L$. Without loss of generality, X is a set in \mathbb{M} : if not then it is isomorphic to one, since every subgroup in \mathcal{F} stabilizes some set, and the action $G \curvearrowright \mathbb{M}$ transfers across the isomorphism.

Consider X as a set of variables and let $\Sigma \subseteq \mathcal{L}(X)$ be the set of sentences of the form

$$\xi(F, K, L, h) = \bigvee_{x \in \bar{F}} \neg hx, \quad \eta(F, K, L, h) = \bigvee_{x \in \bar{F}} hx$$

for $X_{\langle F, K, L \rangle}$ as above and $h \in H$. Then $\Sigma \in \mathbb{M}$.

Lemma 6.3 *If Σ is consistent (in \mathbb{M}) then \mathcal{F} has the Ramsey property.*

Proof Let $v : X \rightarrow \{\top, \perp\}$ be a model of X in \mathbb{M} , and let $K \in \mathcal{F}$ stabilize v . We may assume that $K \leq H$ by replacing K by $K \cap H$ if necessary, since $K \cap H$ certainly stabilizes v if K does. Let $L \in \mathcal{F}$ be a subgroup with $L \leq H$; we show that L is a Ramsey subgroup of H .

Suppose not, and let L be a counterexample. Let $F \subseteq K/L \subseteq H/L$ be a finite subset, such that each finite subset $Y \subseteq F$ has a partition $Y = Y_\top \cup Y_\perp$ with $gF \not\subseteq Y_\top, Y_\perp$ for all $g \in G$. Then $X_{\langle F, K, L \rangle} \in X$ and $\xi(F, K, L, 1), \eta(F, K, L, 1) \in \Sigma$. Hence v is not constant on \bar{F} . But all the variables in \bar{F} lie in the same K -orbit, so this is impossible since v is fixed by K .

This is the required contradiction. □

We now prove a very suggestive result, bearing in mind that we can't assume the sentential compactness theorem and therefore can't prove that Σ is consistent.

Lemma 6.4 *Every finite subset of Σ has a model in \mathbb{M} .*

Proof Let $\Sigma_0 \subseteq \Sigma$ be finite. Each sentence in Σ involves only one element from a single H -translation hT of the K -orbit T , so we may assume without

loss of generality that Σ_0 consists entirely of sentences all of whose variables come from such hT . In fact, we may take $h = 1$. We can then recover Σ_0 by left-multiplication by h , and piece together the valuations (in a sense made precise) to obtain a model for a general finite subset of Σ .

Let $Y = \{x \in T : x \text{ appears in } \Sigma_0\}$. By definition, there exists a partition $Y = Y_\top \cup Y_\perp$ of Y with $hF \not\subseteq Y_\top$ and $hF \not\subseteq Y_\perp$. Define a valuation $v : X \rightarrow \{\top, \perp\}$ by

$$v(x) = \begin{cases} \top & \text{if } x \in Y_\top \\ \perp & \text{if } x \in Y_\perp \\ * & \text{if } x \notin Y \end{cases}$$

where $*$ denotes an arbitrary value which does not matter. This allows us to piece together the valuations for the general case.

Now, v defines a model of Σ_0 , and by the above reasoning it follows that every finite subset of Σ has a model. \square

Theorem 6.5 *Suppose all atoms are \mathcal{F} -symmetric and each $H \in \mathcal{F}$ occurs as a stabilizer of some set in \mathbb{M} . Then the following are equivalent:*

- (i) $\mathbb{M} \models \text{BPIT}$;
- (ii) \mathcal{F} is a Ramsey filter of subgroups of G .

Moreover, $\mathbb{M} \models \text{AC}$ if and only if $\{1\} \in \mathcal{F}$.

Proof By Theorem 2.4, it suffices to prove

$$\text{SCT holds in } \mathbb{M} \xrightarrow{(a)} \mathcal{F} \text{ is a Ramsey filter} \xrightarrow{(b)} \text{AMIT holds in } \mathbb{M}$$

We use all notation as above.

- (a) Suppose $\mathbb{M} \models \text{SCT}$; then Σ has a model by Lemma 6.4. By Lemma 6.3, for an arbitrary $H \in \mathcal{F}$ we can find $K \leq H$ such that every subgroup of K is a Ramsey subgroup. But then the set \mathcal{M} of such K 's for each $H \in \mathcal{F}$ forms a basis for \mathcal{F} , so that \mathcal{F} has the Ramsey property.
- (b) Conversely, suppose that \mathcal{F} is a Ramsey filter, and let \mathcal{B} be the basis of \mathcal{F} consisting of those $K \in \mathcal{F}$ arising in the manner described above. Fix a (nontrivial) distributive lattice \mathbb{L} in \mathbb{M} , and find some $K \in \mathcal{B}$ stabilizing \mathbb{L} .

Using the axiom of choice in the meta-theory, we can find a maximal ideal $I \subseteq \mathbb{L}$ with $K \leq \text{stab}_G(I)$. Then I is \mathcal{F} -symmetric, so lies in \mathbb{M} , and by maximality $\widehat{I} = I$. It remains to prove that I is prime.

Suppose not. Then there exist $x, y \in \mathbb{L} - I$ with $x \wedge y \in I$. Then the ideal generated by

$$I \cup \{kx : k \in K\}$$

is K -invariant, so lies in \mathbb{M} , and properly contains I . Replacing x by y yields the same result for $I \cup \{ky : k \in K\}$. Hence there exist $k_1, \dots, k_r \in K$ with

$$k_1x \vee \dots \vee k_rx \in I^* \quad \text{and} \quad k_1y \vee \dots \vee k_ry \in I^*$$

Let $L = \text{stab}_K(\langle x, y \rangle)$. Then $L \in \mathcal{F}$ and $L \leq K$, and since \mathcal{F} is a Ramsey subgroup, $X = \text{orb}_K(\langle x, y \rangle) \cong K/H$ has the Ramsey property. Let $F = \{k_1\langle x, y \rangle, \dots, k_r\langle x, y \rangle\}$ and from the Ramsey property obtain a finite set

$$Y = \{y_1\langle x, y \rangle, \dots, y_s\langle x, y \rangle\}$$

where $y_1, \dots, y_s \in K$ are such that whenever Y is partitioned into two pieces, one of the pieces contains kF for some $k \in K$.

For $C \subseteq \{1, 2, \dots, s\}$ let $z_C = \left(\bigwedge_{i \in C} y_i x \vee \bigwedge_{j \notin C} y_j x \right)$. Then

$$z = \bigwedge_{C \subseteq \{1, 2, \dots, s\}} z_C = \bigvee_{i=1}^s (y_i x \vee y_i y) = \bigvee_{i=1}^s y_i \cdot x \wedge y \in I$$

Fix a subset $C \subseteq \{1, 2, \dots, s\}$. This determines a partition $Y = Y_\top \cup Y_\perp$, where

$$Y_\top = \{y_i\langle x, y \rangle : i \in C\} \quad \text{and} \quad Y_\perp = \{y_j\langle x, y \rangle : j \notin C\}$$

By choice of Y , there exists $k \in K$ with either $kF \subseteq Y_\top$, in which case,

$$z_C \geq \bigvee_{i \in C} y_i x \geq \bigvee_{j=1}^r k k_j x \in k I^* = I^*$$

or $kF \subseteq Y_\perp$, in which case

$$z_C \geq \bigvee_{i \notin C} y_i x \geq \bigvee_{j=1}^r k k_j y \in k I^* = I^*$$

But then $z = \bigwedge_{C \subseteq \{1, \dots, s\}} z_C \in I^*$. But we saw above that $z \in I$, so $I \cap I^* = \emptyset$, contradicting Lemma 2.9.

We thus have (a) \leftrightarrow (b). Now if $\{1\} \in \mathcal{F}$ then $\text{fix}_G(x) \in \mathcal{F}$ for each $x \in \mathbb{M}$ by upwards-closure, so every set in \mathbb{M} has a wellorder in \mathbb{M} by Lemma 3.8. Conversely, if every set can be wellordered then $\text{fix}_G(A) = \{1\} \in \mathcal{F}$. By Lemma 1.2, the axiom of choice holds in \mathbb{M} if and only if $\{1\} \in \mathcal{F}$. \square

6.1 Post mortem

The ‘caveat’

As alluded to in its statement, the correspondence in Theorem 6.1 is not quite exact. Both the hypotheses and the specification of the normal filter \mathcal{F} differ between Theorems 6.2 and 6.5. The hypotheses differ in that:

- Theorem 6.2 requires that G be Hausdorff and that it have small open subgroups;
- Theorem 6.5 requires that all atoms be symmetric and that each subgroup in the filter \mathcal{F} stabilise some set in \mathbb{M} .

Fortunately this has no implications for our correspondence. The requirement that G be Hausdorff corresponds exactly with our requirement that all atoms be symmetric. If it weren’t Hausdorff then Propositions 5.2 and 5.3 give that some quotient G/N is Hausdorff—and is extremely amenable if and only if G is—and Proposition 3.6 tells us that the model \mathbb{M} is unaffected by this modification. Proposition 3.4 takes care of the requirement that each subgroup in \mathcal{F} stabilise some set in the model.

The specification of the normal filter \mathcal{F} , however, *does* affect the exactness of the correspondence. Specifically:

- In Theorem 6.2, the Ramsey property of \mathcal{F} must be witnessed by G ;
- In Theorem 6.5, the Ramsey property of \mathcal{F} can be witnessed by any subgroup of G .

By Proposition 3.5, if the Ramsey property in Theorem 6.5 is witnessed by $H \leq G$ then we can replace \mathcal{F} by its shrink \mathcal{F}_H without changing the model, and then Theorem 6.2 tells us that H is extremely amenable. Note that this tells us nothing about the extreme amenability or otherwise of G .

Mostowski’s ordered model revisited

Recall that in Mostowski’s ordered model the set A of atoms was equipped with a linear order $<_A$ making $\langle A, <_A \rangle$ order-isomorphic to $\langle \mathbb{Q}, < \rangle$, a group G of all $<_A$ -automorphisms of $\langle A, <_A \rangle$, and a normal filter \mathcal{F} generated by the fixators of finite subsets of A . We endow G with the product topology inherited from A^A in the obvious way, which in turn is given the product topology from the discrete topology on A .

Lemma 6.6 $G \cong \text{Aut}(\mathbb{Q}, <) has small open subgroups.$

Proof Define maps $\pi_a : G \rightarrow A$ for $a \in A$ by $\pi_a(f) = f(a)$. The discrete topology on A is generated by its singletons, so since inverse images preserve unions and intersections, the product topology \mathcal{O} on G is generated by finite intersections of sets of the form $\pi_a^{\leftarrow}(\{b\})$ for $a, b \in A$.

If $1 \in \pi_a^{\leftarrow}(\{b\})$ then $a = b$ and $\pi_a^{\leftarrow}(\{b\}) = \text{fix}_G(\{a\})$. For a finite subset $F \subseteq A$,

$$\bigcap_{a \in F} \text{fix}_G(\{a\}) = \text{fix}_G(F) \leq G$$

and so the fixators of finite subsets of A are a neighbourhood base for 1. In particular, if $U \in \mathcal{O}$ then $\text{fix}_G(F) \subseteq U$ for some finite $F \subseteq A$. \square

Theorem 6.7 $G \cong \text{Aut}(\mathbb{Q}, <)$ is extremely amenable.

Proof By Lemma 6.6, G has small open subgroups, and certainly every atom is symmetric and every subgroup of G lying in \mathcal{F} stabilizes some set in \mathbb{M} . By Theorems 4.1 and 4.7, $\mathbb{M} \not\models \text{AC}$ and $\mathbb{M} \models \text{BPIT}$, so we may apply Theorem 6.5 to get that \mathcal{F} is a Ramsey filter of subgroups of G and $\{1\} \notin \mathcal{F}$.

Working through the proof of Theorem 6.5 with $H = G$ we see that $K = G$ works as a witness to the Ramsey property on \mathcal{F} , so that by Theorem 6.2 G is extremely amenable. \square

A direct proof that $\text{Aut}(\mathbb{Q}, <)$ is extremely amenable can be found in [17].

Further research

Andreas Blass provides some open questions in [4] concerning Ramsey filters; for instance, if G is a group, H is a Ramsey subgroup of G and K is a Ramsey subgroup of H , is K a Ramsey subgroup of G ?

Other avenues of thought include the following:

- Is there an interesting Galois theory surrounding Proposition 3.7?
- What properties do the models of $\text{ZFA} + (\neg\text{AC}) + \text{BPIT}$ corresponding with other groups have?
- Given a formula ϕ implied by the axiom of choice, is there a class of groups satisfying a certain condition that correspond with models of $\text{ZFA} + (\neg\text{AC}) + \phi$?

References

- [1] Bernhard Banaschewski. Prime elements from prime ideals. *Order*, 2:211–213, 1985.
- [2] Andreas Blass. Injectivity, projectivity, and the axiom of choice. *Transactions of the American Mathematical Society*, 255:31–59, 1979.
- [3] Andreas Blass. Prime ideals yield almost maximal ideals. *Fundamenta Mathematicae*, 127(1):57–66, 1987.
- [4] Andreas Blass. Partitions and permutation groups. *Contemporary Mathematics*, 558:453–466, 2011.
- [5] Andreas Blass. Models of ZFA corresponding exactly with a particular class of groups. MathOverflow, 2013. <http://mathoverflow.net/questions/119673> (version: 2013-01-23).
- [6] T. Forster. *Logic, Induction and Sets*. London Mathematical Society Student Texts. Cambridge University Press, 2003.
- [7] Lorenz J. Halbeisen. *Combinatorial Set Theory With a Gentle Introduction to Forcing*. Springer Monographs in Mathematics. Springer, 2011.
- [8] Wojciech Herer and Jens Peter Reus Christensen. On the existence of pathological submeasures and the construction of exotic topological groups. *Mathematische Annalen*, 213:203–210, 1975.
- [9] Thomas J. Jech. Fraenkel–Mostowski models. In *Lectures in Set Theory with Particular Emphasis on the Method of Forcing*, volume 217 of *Lecture Notes in Mathematics*, pages 122–125. Springer Berlin Heidelberg, 1971.
- [10] Thomas J. Jech. *The Axiom of Choice*. Dover, 2008.
- [11] Peter T. Johnstone. Almost maximal ideals. *Fundamenta Mathematicae*, (123):197–209, 1984.
- [12] Peter T. Johnstone. *Notes on Logic and Set Theory*. Cambridge Mathematical Textbooks. Cambridge University Press, 1987.
- [13] Alexander Kechris, Vladimir Pestov, and Stevo Todorćević. Fraïssé limits, Ramsey theory, and topological dynamics of automorphism groups. *Geometric and Functional Analysis*, 15:106–189, 2005.
- [14] Ig Sung Kim. A study on the Boolean algebras in the Fraenkel–Mostowski topos. *Journal of the Korean Society of Mathematical Education. Series B. The Pure and Applied Mathematics.*, 7(1):19–26, 2000.

- [15] Philipp Kleppmann. Duals of filters and ideals. Unpublished manuscript, 2013.
- [16] Theodore Mitchell. Topological semigroups and fixed points. *Ill. J. Math.*, 14:630–641, 1970.
- [17] Vladimir G Pestov. On free actions, minimal flows, and a problem by Ellis. *Transactions of the American Mathematical Society*, 350(10):4149–4166, 1998.
- [18] Andrew M. Pitts. Nominal sets and their applications. Lectures at Midland Graduate School, University of Nottingham, 2011.
- [19] John K. Truss. Permutations and the axiom of choice. In R. Kaye and D. Macpherson, editors, *Automorphisms of First-Order Structures*, pages 131–152. Clarendon Press, 1994.