Symmetric sets and graph models of set and multiset theories

Hoang-Vu Dang
Trinity College

September 8, 2010

This dissertation is submitted for the degree of Doctor of Philosophy

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

Contents

0.1 Summary

This thesis examines the interaction between two ideas. The first is the concept of hereditarily symmetric sets introduced in [Forster 1] as a model of stratified fragment of Zermelo-Fraenkel (ZF) set theory in which the Axiom of Choice fails, thus providing a middle ground between ZF and Quine's set theory New Foundations (NF). We generalise Forster's concept of hereditarily symmetric sets to define a family of models for stratified fragments of ZF, all refuting Choice. These models provide an insight to a question in [Forster 1] about the sensitivity of the hereditarily symmetric sets to changes in their definition. As an application we also include a result by Zachiri McKenzie where these models are used to distinguish two different structures in [Forster 1], namely the hereditarily symmetric sets and the stratified constructible sets.

The second idea, which appeared in various forms in [Hinnion 1], [Holmes 1] and an unpublished paper of Thomas Forster, is to use well-founded extensional relations to create syntactic models of set theory. Here we will introduce a closely related technique to create syntactic models of both ZF and ZFA (where the Axiom of Foundation is replaced with its opposite, the Axiom of anti-foundation) from models of the stratified fragment of ZF. This technique uses graphs in an existing model of the stratified set theory to represent the membership relation of sets in a new model, and the properties of the constructed model depend on the class of graphs we considered initially. In particular restricting those graph to well-founded extensional relations as in the accounts above give us a well-founded construction, while allowing graphs with loops gives us a model of anti-foundation. This construction is applicable to the family of generalised hereditarily symmetric models described previously, and shows that the apparently weak theory of these models is strong enough to interpret ZF. In fact if the Axiom of Choice holds in the original ZF model (from which the generalised hereditarily symmetric models are defined), we can recover an exact replica of this original model within the generalised hereditarily symmetric models.

Finally we utilise both ideas in the context of multisets. We formalise the language and axioms for a multiset theory in which multiplicities are of the same type as sets, and define the multiset analogue of the hereditarily symmetric sets in this theory. Using an adaptation of our previous graph-based technique, we prove the consistency of our multiset theory with anti-foundation and show that unlike in set theory, the subset relation on multisets under our definition need not be antisymmetric; hence an extra axiom is necessary if one wants to ensure the antisymmetry of the inclusion relation.

0.2 Syntactic conventions

To keep the formulae readable we shall keep the number of brackets to a minimum and assign descending priorities to the following symbols:

- Universal and existential quantifiers and their variants.
- Conjunction and disjunction.
- Implications.

For example the formula $P \Rightarrow (\forall x)Q(x) \vee R$ associates as $P \Rightarrow \Big(\big((\forall x)Q(x)\big) \vee R\Big)$.

If R is a binary relation or binary predicate symbol, write

$$(\forall xRy)\phi(x) \Leftrightarrow_{df} (\forall x)(xRy \Rightarrow \phi(x))$$

and

$$(\exists xRy)\phi(x) \Leftrightarrow_{df} (\exists x)(xRy \wedge \phi(x))$$

For multiple quantifiers we write

$$(\forall x_1 \ldots x_n)\phi(x_1 \ldots x_n) \Leftrightarrow_{df} (\forall x_1) \ldots (\forall x_n)\phi(x_1 \ldots x_n)$$

and

$$(\exists x_1 \ldots x_n)\phi(x_1 \ldots x_n) \Leftrightarrow_{df} (\exists x_1) \ldots (\exists x_n)\phi(x_1 \ldots x_n)$$

Similarly we extend this notation to

$$(\forall \langle x_1 \ldots x_n \rangle Ry)\phi(x_1 \ldots x_n) \Leftrightarrow_{df} (\forall x_1 \ldots x_n)(\langle x_1 \ldots x_n \rangle Ry \Rightarrow \phi(x_1 \ldots x_n))$$

and

$$(\exists \langle x_1 \ldots x_n \rangle Ry)\phi(x_1 \ldots x_n) \Leftrightarrow_{df} (\exists x_1 \ldots x_n)(\langle x_1 \ldots x_n \rangle Ry \wedge \phi(x_1 \ldots x_n))$$

0.3 Other notations

In this thesis relations and functions mean sets of ordered pairs (or tuples in the case of higher arity), unless specified otherwise.

Write $f \in \mathit{Function}$ to mean f is a function and f^{-1} for the inverse $\{\langle y, x \rangle : \langle x, y \rangle \in f\}$ (which may not be a function).

Write $\operatorname{dom} f$ for the domain of the function f and $\operatorname{ran} f$ for the range of f. If R is a relation, write $\operatorname{Dom} R$ for the field of R, i.e. the set $\{x : (\exists y)(\langle x, y \rangle \in R \vee \langle y, x \rangle \in R)\}$.

Remark 1. *If f is a function, then it is also a relation. Thus we have*

$$\mathrm{Dom}\, f = \mathrm{dom}\, f \cup \mathrm{ran}\, f$$

Write $TC(x)$ for the set-theoretic transitive closure of x, i.e. the smallest transitive set that contains x.

Write $Trans(R)$ for the transitive closure of the relation R, i.e. the smallest transitive relation that includes R.

Definition 1. *ι is the function-class $x \mapsto \{x\}$.*

Definition 2. *j is the operator taking any function-class f to jf, where $(jf)x := \{f(y) : y \in x\}$.*

Remark 2. *Note that by extending f with the identity function where necessary, we can define jf on the whole universe of sets; thus following this convention jf will always be a proper class. However in practice we can always restrict our attention to some set as domain, where the size of jf will not be a problem.*

Following tradition we write $f``x$ as a shorthand for $(jf)x$, so for example $\iota^2``x = \{\{\{y\}\} : y \in x\}$.

If our theory is strong enough to implement natural numbers, as is often the case, we can formally implement j^n for any natural number n of the theory as follows:

Definition 3. *Write*

$$\begin{aligned} y = j^n f(x) \Leftrightarrow_{df} (\exists s \in Function)\Big(& \mathrm{dom}\, s = \{m \in \mathbb{N} : m \leq n\} \wedge \\ & \forall i > 0)\phi(s(i), s(i+1)) \wedge s(n)(x) = y \wedge \\ & f \subset s(0) \wedge f^{-1} \subset s(0) \wedge \bigcup^n x \subset \mathrm{dom}\, s(0) \wedge \\ & (\forall z \in \mathrm{dom}\, s(0) \setminus (\mathrm{dom}\, f \cup \mathrm{ran}\, f))s(0)(z) = z\Big) \end{aligned}$$

where $\mathbb{N}$ is the set of natural numbers within the theory and $\phi(f,g)$ is shorthand for

$$\begin{aligned} & f \in Function \wedge g \in Function \wedge \mathrm{dom}\, g = \mathcal{P}\, \mathrm{dom}\, f \wedge \\ & (\forall x \in \mathrm{dom}\, g) g(x) = \{f(y) : y \in x\} \end{aligned}$$

Remark 3. *Note that in the formula above, the function $s(0)$ is an extension of f by the identity function to a larger domain. This ensures that x will be in the domain of $s(n)$.*

When we refer to a permutation with no specified domain, we mean just a bijection whose domain and range are the same.

A formula ϕ is *stratified* if we can assign to each variable x in ϕ a natural number $\mathbf{t}(x)$, the type of x, such that $\mathbf{t}(x) = \mathbf{t}(y)$ whenever $x = y$ is a subformula of ϕ, and $\mathbf{t}(y) = \mathbf{t}(x) + \mathbf{1}$ whenever $x \in y$ is a subformula of ϕ.

A set x is *Cantorian* if it is the same size as the set of singletons of its members; it is *strongly Cantorian* if the singleton operator is the bijection that makes the set Cantorian.

The theory strZF consists of stratified instances of ZF axioms: Extensionality, stratified instances of Comprehension, Pairing, Union, Power Set, stratified instances of Collection, Foundation, and Infinity (there exists a well-order with no maximal element).

Remark 4. *Without full Comprehension, there is a difference between Collection and Replacement - here we choose Collection to suit the constructions we are going to use.*

New Foundations (NF) is the theory consisting of Extensionality, Comprehension and the existence of a universal set

$$(\exists V)(\forall x)(x \in V)$$

NFU is a variant of NF where Extensionality is restricted to non-empty sets, and has been proven to be consistent with ZF.

We denote the class of all ordinals by ON. In ZF models this means the class of von Neumann ordinals.

0.3.1 Symbol overloading

To avoid too complicated notations, we will overload certain symbols to denote analogous concepts in different contexts. For example, $[A, a]$ might denote a pointed digraph when we build a syntactic model of set theory, or a pointed hypergraph when we intend to provide an interpretation of multiset theory. Similarly $f^{-1}(x)$ may be the preimage of x under the function of f, while $R^{-1}(x)$ is the set of predecessors of x under the relation R. In such circumstances, the meaning of the overloaded symbol will be clear from the specified context.

1 HS and its variants

The structures presented in this section have their roots in the class of hereditarily symmetric sets HS, presented by Thomas Forster in [Forster 1]. Here we will give a very brief discussion of this structure, leading to a few questions that motivate the construction in the next section.

Definition 4. *(Forster) Let G be the group of finite permutations on V_ω, or the alternating group on V_ω. For $n \in \omega$, say x is* n-symmetric *if $(\forall i \geq n)(\forall \sigma \in G) j^i\sigma(x) = x$. Say x is* symmetric *if x is n-symmetric for some n.*

HS is the class of all hereditarily symmetric sets*, i.e. $x \in HS$ if and only if x and everything in its transitive closure is symmetric.*

Remark 5. *While in [Forster 1] n is a concrete natural number, here we will take n to be any finite ordinal in V in order for HS to be a definable class.*

Consider the cumulative hierarchy $\{V_\alpha : \alpha \in ON\}$. It is easy to see that any set of rank n is $n+1$-symmetric for all $n \in \omega$, hence $V_\omega \subset HS$. There is also a good reason for letting G act on V_ω only: Suppose G acts transitively on V_α for some $\alpha > \omega$, and $x \notin V_\omega$ is symmetric. Then $TC(x)$ must contain the whole of V_α, hence $HS = V_\omega$!

In [Forster 1], HS is shown to be a model of strZF. Furthermore $V_\alpha \cap HS \in HS$ for all ordinals α. We will omit the proof here, since a generalised result will be shown in the next section. However since in the definition of symmetry we take n to be finite ordinals instead of concrete natural numbers, the central device in the proof needs a slight modification. The following lemma is adapted from [Coret 1]:

Lemma 1. *(Coret's Lemma) Let $\phi(x_1 \ldots x_k)$ be a stratified formula such that x_i have type $\mathbf{t}_i$ in a stratification of ϕ and let t_i be the von Neumann ordinal corresponding to $\mathbf{t}_i$.*

Then for any permutation $\sigma \in V$

$$(\forall m \in \omega)\Big(\phi(x_1 \ldots x_k) \Leftrightarrow \phi(j^{t_1+m}\sigma(x_1) \ldots j^{t_k+m}\sigma(x_k))\Big)$$

Proof. Induction on the structure of ϕ.

The original Coret's Lemma was proved for concrete natural numbers as types, whereas here we have converted all the types into von Neumann ordinals. However with our formal definition of j^n (Definition 3) the lemma can be proved in exactly the same way as in the original lemma. $\square$

Remark 6. *From now on, whenever we refer to the types of variables in a stratification, we mean the von Neumann counterparts to the concrete natural types. In other words, a stratification in our context is a map from the set of variables to the finite von Neumann ordinals.*

In that sense, the notion of stratification technically depends on the model V. However in practical terms it does not matter, since the finite ordinals used as types are always standard. By using finite ordinals in V instead of concrete natural numbers, we will be able to apply Coret's Lemma to our formal definition of hereditarily symmetric sets above.

One peculiar feature of HS is that the Axiom of Choice fails in it, regardless of the status of Choice in V: there is no total ordering of V_ω in HS. To see this, suppose R is an n-symmetric total ordering of V_ω and let σ be the transposition swapping $\emptyset$ and $\{\emptyset\}$. Since $j^{n+3}\sigma(R) = R$, $\langle \iota^n\emptyset, \iota^{n+1}\emptyset\rangle \in R$ if and only if $\langle \iota^{n+1}\emptyset, \iota^n\emptyset\rangle \in R$. This means R cannot be antisymmetric.

This bears some similarity with the theory of New Foundations (NF), which consists of Extensionality, a Universal set and stratified Comprehension, and disproves Choice. Thus it was hoped that HS might provide a model for some intermediate theory with NF-like properties but consistent with ZF. In particular, [Forster 1] questioned the existence of an initial segment $HS \cap V_\alpha$ with a set containing all isomorphism types of well-orders (in said initial segment).

There is a list of stratified rudimentary functions in [Forster 1], which are all absolute for transitive structures, such that any set closed under those and power set is closed under stratified Δ_0 comprehension. Using these, we can build a structure S along the lines of Gödel's L:

- $S_0 := \emptyset$
- $S_{\alpha+1} :=$ closure of $S_\alpha \cup \{S_\alpha\}$ under stratified rudimentary functions.
- $S_\lambda := \bigcup_{\alpha<\lambda} S_\alpha$ for limit λ
- $S := \bigcup_{ON} S_\alpha$

The class S is also a model of strZF, and has an external well-ordering just like L. In fact it is easy to show that $S \subset L$: the stratified rudimentary functions are absolute and L is a model of ZF itself, so there exists a version of S inside L which is the same as S in V. However, $S \subset HS$ and $V_\omega \in S$, which means Choice also fails in S. The natural question to ask here is if $S = HS \cap L$?

Remark 7. *With our earlier choice of the group of permutation G, it is easy to show that $G \in L$. Thus we can construct HS relative to L, and this new class turns out to be the same as $HS \cap L$.*

Instead of looking into HS itself, we will answer this by looking at a family of structures related to HS. These structures also provides an insight into the question in [Forster 1] about the sensitivity of the hereditarily symmetric sets to the group of permutations that is used in the definition of symmetry.

There is also an admittedly vague point of discussion about how much information is lost when we move from V to HS. By requiring the set to be symmetric, we have seemingly destroyed much of the unstratified information, leading to the loss of Choice and all infinite von Neumann ordinals. It may be a little surprising that we can in some sense recover all the information about V just by looking inside HS if Choice originally holds, or at least recover a model of ZF if not.

1.1 The construction

We will start by generalising the definition of symmetric sets. The new definition parallels Fraenkel-Mostowski-Specker (FMS) models as described in [Felgner 1], but we are only interested in preserving stratified Comprehension. Symmetric sets in the FMS method are those preserved under $\in$-isomorphisms of V, so to get non-trivial models one would need to start with non-trivial isomorphisms, i.e. ZF with atoms. Underlying our construction is the concept of a stratimorphism, i.e. an ω-collection of permutations f_i such that $x \in y \Leftrightarrow f_i(x) \in f_{i+1}(y)$; in other words $f_{i+1} = jf_i$. By using stratimorphisms in place of the stronger $\in$-isomorphisms, we no longer need to assume the existence of atoms in our initial model.

Remark 8. *For the sake of simplicity we will not formally define stratimorphisms but make direct changes to the definition of our permutation models. The result is essentially the same as if one simply replaced the $\in$-isomorphisms in the FMS construction with stratimorphisms.*

We work in a model V of ZF. Let G be a group of permutations (i.e. bijections f where $\operatorname{dom} f = \operatorname{ran} f$) and $\mathcal{F}$ a filter on G.

Definition 5. *Say a permutation σ* fixes *x* above *$n \in \omega$ if and only if $j^i\sigma(x) = x$ for all $i \geq n$.*

Definition 6. *The n*-stabiliser *of x in G, $G_n(x)$, is the set of $\sigma \in G$ that fixes x above n.*

The weak stabiliser *of x in G, $G_\omega(x)$, is the set of $\sigma \in G$ that fixes x above some $n \in \omega$. In other words $G_\omega(x) = \bigcup_{n\in\omega} G_n(x)$.*

Definition 7. *$x \in V$ is* strongly symmetric *(with respect to G and $\mathcal{F}$) if $G_n(x) \in \mathcal{F}$ for some $n \in \omega$.*

$x \in V$ is weakly symmetric *(with respect to G and $\mathcal{F}$) if the $G_\omega(x) \in \mathcal{F}$.*

Clearly a set is weakly symmetric if it is strongly symmetric, since $\mathcal{F}$ is closed under inclusion.

Let HSS and HWS be the classes of hereditarily strongly symmetric sets and hereditarily weakly symmetric sets respectively. They are definable in V as follows:

Define

$$G_n(x) := \{f \in G : (\forall m \geq n) x = j^m f(x)\}$$

Similarly

$$G_\omega(x) := \{f \in G : (\exists n \in \omega)(\forall m \geq n) x = j^m f(x))$$

Thus we can define

$$x \in HSS \Leftrightarrow_{df} (\forall y \in TC\{x\})(\exists n \in \omega)(G_n(y) \in \mathcal{F})$$

$$x \in HWS \Leftrightarrow_{df} (\forall y \in TC\{x\})(G_\omega(y) \in \mathcal{F})$$

Remark 9. *The class HS in [Forster 1] thus corresponds to HSS when G is the group of finite permutations of V_ω (or the alternating group) and $\mathcal{F} = \{G\}$.*

For these classes to be models of strZF, we will need to impose extra conditions later on.

1.2 Axioms of strZF

Lemma 2. *Let $\phi(x)$ be a stratified formula with all parameters strongly (respectively weakly) symmetric. If $(\exists! x)\phi(x)$, then that unique witness x is strongly (respectively weakly) symmetric.*

Proof. Let $a_1 \ldots a_n$ be the parameters of $\phi(x)$, where a_k has type $t_k \geq 0$ and x has type $t \geq 0$ in some stratification of ϕ. Suppose $\phi(x, a_1 \ldots a_n)$ holds.

If the a_k are strongly symmetric and $G_{m(k)}(a_k) \in \mathcal{F}$, then $H := \bigcap_k G_{m(k)}(a_k) \in \mathcal{F}$.

Let $m := \max\{m(1) \ldots m(n)\}$. By Coret's Lemma, for any $i \geq 0$ and $\sigma \in H$

$$\phi(j^{t+m+i}\sigma(x), j^{t_1+m+i}\sigma(a_1) \ldots j^{t_n+m+i}\sigma(a_n))$$

But this is exactly $\phi(j^{t+m+i}\sigma(x), a_1 \ldots a_n)$, so by uniqueness of x we have $x = j^{t+m+i}\sigma(x)$. Hence $H \subset G_{t+m}(x)$, and thus x is strongly symmetric.

If the a_k are weakly symmetric, take

$$H := \bigcap_k G_\omega(a_k) \in \mathcal{F}$$

If $\sigma \in H$, let $\sigma \in G_{m(i)}(a_i)$ for each a_i and $m := \max\{m(1) \ldots m(n)\}$. By the same argument as above, $x = j^{t+m+i}\sigma(x)$ for any $i \geq 0$, so $H \subset G_\omega(x)$. □

Lemma 3. *HSS and HWS safisfy Extensionality.*

Proof. The two classes are transitive. □

Lemma 4. *HSS and HWS satisfy Empty Set.*

Proof. Easy to see $G_1(\emptyset) = G \in \mathcal{F}$ so $\emptyset \in HSS \subset HWS$. □

Lemma 5. *HSS and HWS satisfy Pair Set.*

Proof. Let $x, y \in HSS$. By Lemma 2 $\{x, y\}$ is strongly symmetric so it is in HSS.

Similarly for HWS. □

Lemma 6. *HSS and HWS satisfy Sumset.*

Proof. Let $x \in HSS$. By Lemma 2, $\bigcup x$ is strongly symmetric so it is in HSS.

Similarly for HWS. □

Lemma 7. *If $x, y \in HSS$, then $x \times y \in HSS$, and similarly for HWS.*

Proof. If $x, y \in HSS$, then $x \times y$ is strongly symmetric by Lemma 2. With Pair Set it is easy to see that $x \times y \in HSS$.

Similarly for HWS. □

Lemma 8. *Let $\phi(x)$ be a stratified formula with parameters in HSS and $A \in HSS$. Then $\{x \in A : \phi(x)\} \in HSS$.*

Similarly for HWS.

Proof. Let A and the parameters of ϕ be in HSS. The formula $y = \{x \in A : \phi(x)\}$ is stratified and y is unique given A and the parameters of ϕ. By Lemma 2 the set $\{x \in A : \phi(x)\}$ is strongly symmetric and thus in HSS as it is a subset of A.

Similarly for HWS. □

Corollary 1. *HSS and HWS satisfy stratified Δ_0 Comprehension.*

Proof. By the last lemma and absoluteness of Δ_0 formulae. □

Although $y = \mathcal{P}(x)$ is stratified, we cannot prove Power Set in the same way as for the last few lemmata. A similar proof merely shows that $\mathcal{P}(x)$ is symmetric. For example, if G is the group of permutations which are themselves members of V_ω, and $\mathcal{F} = \{G\}$, then $V_\omega \in HSS$ but $V_{\omega+1}$ is not. The power set of x in HSS would be $\mathcal{P}(x) \cap HSS$; but its definition is not stratified since the definition of HSS involves the use of transitive closures.

On the other hand, HSS and HWS contain transitive closures of their members even though the definition of transitive closure is not stratified.

Lemma 9. $(\forall x \in HSS)(TC(x) \in HSS)$ *and similarly for HWS.*

Proof. Let $x \in HSS$ and $G_k(x) \in \mathcal{F}$.

Since $z \in x \to f(z) \in jf(x)$, by an easy induction

$$(\forall f \in Function)(\forall n \in \omega)(z \in \bigcup^n x \to f(z) \in \bigcup^n j^{n+1} f(x))$$

Now let $y \in \bigcup^m x$ and $\sigma \in G_k(x)$. Then for any $n \geq k$, setting $z = y$ and $f = j^n \sigma$ gives

$$j^n \sigma(y) \in \bigcup^m j^{n+1+m} \sigma(x) = \bigcup^m x$$

Thus $j^{k+1}\sigma(TC(x)) \subset TC(x)$ and applying the result with σ^{-1} gives

$$j^{k+1}\sigma(TC(x)) = TC(x)$$

Therefore $G_{k+1}(TC(x)) \supset G_k(x)$, and we already have $TC(x) \subset HSS$.

Let $x \in HWS$ and σ fix x above k. The same argument as above shows that σ fixes $TC(x)$ above $k+1$, so $G_\omega(TC(x)) \supset G_\omega(x)$. □

For HSS and HWS to satisfy Power Set and full stratified Comprehension, we need to impose further conditions on G and $\mathcal{F}$.

Definition 8. *$\langle G, \mathcal{F}\rangle$ satisfies the* strong closure condition *if*

$$G^* := \{\sigma \in G : (\forall n \in \omega)(j^n \sigma``HSS \subset HSS)\} \in \mathcal{F}$$

Similarly $\langle G, \mathcal{F}\rangle$ satisfies the weak closure condition *if*

$$\{\sigma \in G : (\forall n \in \omega)(j^n \sigma``HWS \subset HWS)\} \in \mathcal{F}$$

Remark 10. *If the closure condition holds, a set is strongly symmetric with respect to* $\langle G, \mathcal{F} \rangle$ *if and only if it is strongly symmetric with respect to* $\langle G^*, \mathcal{F} \cap \mathcal{P}G^* \rangle$ *(and the same holds for weak symmetry); so by replacing* G *with* G^* *we can assume that*

$$(\forall n \in \omega)(\forall \sigma \in G)(\forall y \in HSS) j^n\sigma(y) \in HSS$$

(or HWS *respectively).*

Lemma 10. *Assuming the corresponding closure condition,* HSS *and* HWS *satisfy Power set.*

Proof. Let $x \in HSS$ and $G_n(x) \in \mathcal{F}$ where $n \geq 1$.

If $x \supset y \in HSS$ and $\sigma \in G_n(x)$, then $j^n\sigma(y) \subset j^n\sigma(x) = x$ and $j^n\sigma(y) \in HSS$ by the strong closure condition. Hence $G_{n+1}(\mathcal{P}x \cap HSS) \supset G_n(x)$, so $\mathcal{P}x \cap HSS$ is the power set of x in HSS.

Similarly for HWS. □

Lemma 11. *Assuming the corresponding closure condition,* HSS *and* HWS *are proper classes.*

Proof. If $HSS \in V$, then it is strongly symmetric by the strong closure condition. Thus $HSS \in HSS$, which contradicts foundation.

Similarly for HWS. □

Corollary 2. *Assuming the corresponding closure condition,* HSS *and* HWS *have members of every rank in* V*.*

Proof. Since HWS and HSS are proper classes, they have members of arbitrarily large ranks; but they are also transitive. □

Let λ be a limit ordinal with cofinality greater than $\operatorname{rank} G$. We now show that initial segments of HSS and HWS above λ are themselves members of the corresponding structures. With the corollary above in mind, this will give us an unlimited supply of distinct hereditarily symmetric sets.

Lemma 12. $(\forall \alpha \geq \lambda)(\forall n \in \omega)(\forall \sigma \in G)(j^n\sigma``V_\alpha \subset V_\alpha)$.

Proof. We prove the claim by induction on α.

If $\alpha > \lambda$ is a limit: Let $x \in V_\alpha$ and $\lambda < \beta < \alpha$ such that $x \in V_\beta$. Then $j^n\sigma(x) \in V_\beta \subset V_\alpha$.

If $\alpha = \beta + 1$ for some $\beta \geq \lambda$: Let $x \in V_\alpha$.

- If $n \geq 1$, then $j^n\sigma(x) \subset j^n\sigma(V_\beta) \subset V_\beta$ since $x \subset V_\beta$.
- If $n = 0$ the proof is trivial since $\operatorname{ran}\sigma \subset TC(G) \subset V_\lambda$.

If $\alpha = \lambda$: Let $x \in V_\lambda$ and $n \in \omega$, then $\operatorname{rank} j^n\sigma(x) \leq \operatorname{rank} x + \operatorname{rank} G + n$ by induction on n.

- If $n = 0$, then $\operatorname{rank}\sigma(x) < \operatorname{rank} G < \operatorname{rank} x + \operatorname{rank} G$.
- If $n \geq 1$, then

  $$(\forall y \in x)(\operatorname{rank} j^{n-1}\sigma(y) < \operatorname{rank} y + \operatorname{rank} G + (n-1) \leq \operatorname{rank} x + \operatorname{rank} G + (n-1)$$

  so $\operatorname{rank} j^n\sigma(x) \leq \operatorname{rank} x + \operatorname{rank} G + n$. But $\operatorname{rank} x + \operatorname{rank} G + n < \lambda$ since λ has cofinality greater than $\operatorname{rank} G$, so $j^n\sigma(x) \in V_\lambda$.

□

Corollary 3. *Assuming the corresponding closure condition, $HSS \cap V_\alpha \in HSS$ for all $\alpha \geq \lambda$. In particular*

$$(\forall \alpha \geq \lambda)(G_1(HSS \cap V_\alpha) \supset G^*)$$

(respectively for HWS).

Proof. It is enough to prove that $HSS \cap V_\alpha$ is strongly symmetric.

For all $n \geq 1$ and $\sigma \in G^*$, $j^n\sigma(HSS \cap V_\alpha) \subset V_\alpha$ by the lemma, so by the closure condition

$$j^n\sigma(HSS \cap V_\alpha) \subset HSS \cap V_\alpha$$

This also shows $j^n\sigma^{-1}(HSS \cap V_\alpha) \subset HSS \cap V_\alpha$, therefore $j^n\sigma(HSS \cap V_\alpha) = HSS \cap V_\alpha$.

Similarly for HWS. □

This gives us an easy form of Infinity:

Lemma 13. *Assuming the corresponding closure condition, $\exists A \in HSS$ (respectively HWS) such that HSS (respectively HWS) believes A is a well-order with no top element and no limit point.*

Proof. Take a sequence of ordinals $\langle \alpha(n) : n \in \omega \rangle$ all greater than λ, and let

$$A := \{HSS \cap V_{\alpha(n)} : n \in \omega\}$$

Then $G_2(A) \supset G^*$ so $A \in HSS$, and by Power Set and stratified Δ_0 Comprehension the inclusion ordering on A is in HSS. This is a well-order with no top element and no limit point in V, so it has the same properties in HSS.

Similarly for HWS. $\square$

We can also prove stratified Comprehension and Collection now.

Lemma 14. *Assuming the corresponding closure condition, HSS (respectively HWS) satisfies stratified Comprehension.*

Proof. Let $\phi(x)$ be a stratified formula with parameters in HSS and $A \in HSS$. By Reflection in V, let $\alpha > \lambda$ be such that $A \in V_\alpha$ and

$$(\forall x \in V_\alpha)(\phi(x)^{HSS} \Leftrightarrow V_\alpha \models \phi(x)^{HSS})$$

Thus

$$\{x \in A : \phi(x)^{HSS}\} = \{x \in A : \phi(x)^{V_\alpha \cap HSS}\}$$

But $(V_\alpha \cap HSS) \in HSS$ so the set of interest is actually an instance of stratified Δ_0 Comprehension in HSS.

Similarly for HWS. $\square$

Lemma 15. *Assuming the corresponding closure condition, HSS (respectively HWS) satisfies Collection.*

Proof. Let $A \in HSS$ and $\phi(x, y)$ is a formula with parameters in HSS such that

$$(\forall x \in A)(\exists y \in HSS)\phi(x, y)^{HSS}$$

By Collection and Comprehension in V there exists $B \subset HSS$ such that

$$(\forall x \in A)(\exists y \in B)\phi(x, y)^{HSS}$$

Let $\alpha > \operatorname{rank} B$ such that $V_\alpha \cap HSS \in HSS$, then $V_\alpha \cap HSS$ is the required set.

The proof for HWS is identical. $\square$

Corollary 4. *Assuming the corresponding closure condition, HSS (respectively HWS) satisfies stratified Replacement.*

Proof. Stratified Replacement follows from Collection and stratified Comprehension. $\square$

Lemma 16. *If $f \cap HSS$ is a function, then for any $n \in \omega, A \in HSS$, the graph of $j^n f$ on domain A is in HSS; and similarly for HWS.*

Proof. Induction on n.

For $n = 0$, the graph of f on domain A is in HSS by stratified Δ_0 Comprehension from $(A \cup \operatorname{dom} f)^2$. If f is not defined on all of A, we extend it with the identity function.

If $f_n \in HSS$ is the graph of $j^n f$ on domain A, then $f_{n+1} = \{\langle x, y\rangle : (x \in A) \wedge (y = f_n``x)\}$ is strongly symmetric by Lemma 2. To show $f_{n+1} \in HSS$, it is enough to show that $f_n``x \in HSS$ for all $x \in A$. But $f_n``x$ is strongly symmetric also by Lemma 2, and is a subset of $\operatorname{ran} f_n$.

The proof for HWS is the same. $\square$

Corollary 5. *$G \cap HSS \in \mathcal{F}$ and $G \cap HWS \in \mathcal{F}$ each implies the corresponding closure condition.*

Proof. If $\sigma \in G \cap HSS$ and $x \in HSS$, then the graph of $j^n\sigma$ on domain $\{x\}$ is in HSS so we can take $G^* = G \cap HSS$.

Similarly for HWS. $\square$

Remark 11. *Since every set of finite rank in V is in HSS regardless of the choices of G and $\mathcal{F}$, it is clear that the original HS satisfies $G \cap HSS \in \mathcal{F}$. Thus it is a model of strZF.*

Now we turn to answering some of the questions about HS posed in the introduction. The following result by Zachiri McKenzie from [Dang and McKenzie 1] shows that $S \neq HS \cap L$.

Remark 12. *Note that L is a model of ZF, so we can define HS^L and S^L. The stratified rudimentary functions are absolute, so $S^L = S$. The formula "x is symmetric" is also absolute with this particular choice of G and $\mathcal{F}$, so $HS^L = HS \cap L$. Therefore it is enough to show $HS \neq S$ for all models V of ZF.*

Theorem 1. *(Zachiri McKenzie) $HS \neq S$*

Proof. Let HS_1 be the hereditarily strong symmetric sets where G is generated by the single transposition swapping $\{\{V_\omega\}\}$ and $\{\{V_\omega\}, V_\omega\}$, and $\mathcal{F} = \{G\}$.

G has a stratified definition with parameter V_ω, but it is easy to see $V_\omega \in HS_1$, so HS_1 satisfies the strong closure condition. Hence HS_1 is a model of strZF. An induction on α shows $S_\alpha \in HS_1$ for all ordinals α, so $S \subset HS_1$.

Now let $A_0 := \{V_\omega\}$, $A_{n+1} := A_n \cup \{A_n\}$ and $A := \{A_n : n \in \omega\}$. Then $A \in HS \setminus HS_1$ by direct inspection, so $HS \neq S$. $\square$

Remark 13. *Thomas Forster conjectured downward absoluteness from V to HS for the smallest class of sentences that is*

- *Stratified.*
- *Inclusive of all atomic formulae and their negations.*
- *Closed under universal quantifiers.*
- *Closed under unique existential quantifiers, e.g. $(\exists! x)\phi(x)$.*

However, this is not the case if parameters are allowed. For example, let P be the power set of V_ω in HS and consider the sentence

$$(\exists! x)(x = \mathcal{P}(V_\omega) \wedge x \neq P)$$

Clearly this sentence (with V_ω and P as parameters) is in the class defined above, but it is true in V and false in HS since $P \neq V_{\omega+1}$. It is still open if the conjecture holds for sentences without parameters.

Definition 9. *A set $x \in HSS$ is* uniformly (strongly) symmetric *if there are $H \in \mathcal{F}$, $n \in \omega$ such that $(\forall y \in x) H \subset G_n(y)$.*

A set $x \in HWS$ is uniformly (weakly) symmetric *if there exists $H \in \mathcal{F}$ such that $(\forall y \in x) H \subset G_\omega(y)$.*

The following result follows immediately from the definition, and a similar result holds for HWS:

Lemma 17. *Let $x \in HSS$ be uniformly symmetric, then:*

- *$\mathcal{P}(x) \in HSS$ and is uniformly symmetric.*
- *If $R \in V$ is a relation on x, then $R \in HSS$ and is uniformly symmetric.*
- *In particular HSS believes that x is strongly Cantorian.*

Proof. Let $H \in \mathcal{F}$, $n \in \omega$ such that $(\forall y \in x) H \subset G_n(y)$. Then clearly $H \subset G_{n+1}(a)$ for all $a \subset x$ and thus $H \subset G_{n+2}(\mathcal{P}(x))$. Hence $\mathcal{P}(x) \in HSS$ and is uniformly symmetric.

If R is a relation on x, then every $a \in R$ is an ordered pair of things in x, so $H \subset G_{n+2}(a)$. Therefore $H \subset G_{n+3}(R)$ and R is uniformly symmetric.

Finally graph of the singleton function ι on x is a relation on x, so it is hereditarily symmetric i.e. x is strongly Cantorian in the view of HSS. $\square$

Lemma 18. *If Choice holds in V and the strong closure condition holds, then for any limit $\alpha > |G|^+$ there is no $X \in HSS \cap V_\alpha$ such that any well-order in $HSS \cap V_\alpha$ is isomorphic to a member of X.*

Proof. Suppose X satisfies the hypothesis.

If $|TC(X)| > \max\{\omega, |G|\}$, let $G_m(TC(X)) \in \mathcal{F}$ for some m. Let H be the group of permutations generated by

$$\{j^i\sigma : i \geq m, \sigma \in G_m(TC(X))\}$$

For any $x \in TC(X)$ let $\hat{x} := \{\tau(x) : \tau \in H\}$, and let $Y := \{\hat{x} : x \in TC(X)\}$.

Then it is clear that $\hat{x} \subset TC(X)$, so Y partitions $TC(X)$. Assuming Choice in V, the size of each $\hat{x}$ is at most $|H| \leq \max\{\omega, |G|\} < |TC(X)|$ so again by Choice $|Y| = |TC(X)|$.

Furthermore $G_m(X) \subset G_1(\hat{x})$ for all $x \in TC(X)$ by definition of H, so Y is uniformly symmetric. Thus $\mathcal{P}(Y) \in HSS$ and is also uniformly symmetric. Hence the well-ordering of $\mathcal{P}(Y)$ in V is also in HSS, but it is not isomorphic to anything in X since the size of $\mathcal{P}(Y)$ is too large.

If $|TC(X)| \leq \max\{\omega, |G|\}$, there is some $Z \in HSS$ such that $\max\{\omega, |G|\} < \operatorname{rank} Z < \alpha$ since the strong closure condition implies that HSS has members of every rank. Then $\max\{\omega, |G|\} < |TC(Z)|$, so by the same argument as above we can find a well-order in $HSS \cap V_\alpha$ whose carrier set is bigger than $|TC(Z)|$, which therefore cannot be isomorphic to anything in X. $\square$

Remark 14. *In the case of HS we can weaken the condition on α to being any limit ordinal. The reason is that $|G|$ is countable and $V_{\omega+1} \cap HS$ is (externally) uncountable, so the proof above holds for all limit $\alpha > \omega$. The case $\alpha = \omega$ is trivial since all finite von Neumann ordinals are in HS.*

The following lemma is the first step to recovering unstratified information from HSS:

Lemma 19. *If HSS believes that R is a well-founded relation, then R is well-founded in V.*

Similarly for HWS.

Proof. Suppose R is not well-founded in V, so there exists $X \subset \operatorname{Dom} R$ with no R-minimal element.

For HSS, let $G_n(R) \in \mathcal{F}$ and define

$$Y := \{j^i\sigma(x) : x \in X \wedge \sigma \in G_n(R) \wedge i \geq n\}$$

Clearly $G_n(R) \subset G_n(\mathrm{Dom}\, R)$, therefore $Y \subset \mathrm{Dom}\, R$. But $G_n(R) \subset G_{n+4}(Y)$ by definition of Y, so Y is strongly symmetric and thus $Y \in HSS$.

Now let $y := j^i\sigma(x)$ for some $x \in X$, $\sigma \in G_n(R)$, $i \geq n$. Since X has no R-minimal member, there exists $z \in X$ such that $\langle z, x\rangle \in R$. But then

$$\langle j^i\sigma(z), j^i\sigma(x)\rangle = j^{i+2}\sigma\langle z, x\rangle \in R$$

But $j^i\sigma(z) \in Y$, so y is not R-minimal in Y. This shows Y has no R-minimal member, so R is not well-founded in HSS.

Similarly, for HWS we define

$$Y := \bigcup_n \{j^i\sigma(x) : x \in X \wedge \sigma \in G_n(R) \wedge i \geq n\}$$

It is easy to check that $Y \subset \mathrm{Dom}\, R$ and Y is weakly symmetric by definition. As before Y has no R-minimal element so R is not well-founded in HWS. $\square$

Corollary 6. *HSS thinks that R is a well-founded extensional relation on X if and only if V believes the same thing; and similarly for HWS.*

Proof. Extensionality of R is a Δ_0 predicate with parameters X, R and thus is absolute since HSS and HWS are transitive. $\square$

So far we have shown identical results for HSS and HWS. Thus it may be of interest to demonstrate that they are not in fact identical classes.

Remark 15. *Let G be the group of finite permutations of V_ω and $\mathcal{F} = \{G\}$, so that HSS is the same as the class HS in [Forster 1]. Then we can find a set in the corresponding HWS that does not belong to HSS as follows:*

For any $n \in \omega$ let $A_n := \iota^n\{x \in V_\omega : |x| = n\}$. Then for any $\sigma \in G$ and any $i \in \omega$, it is easy to see that $j^i\sigma A_n = A_n$ if $i \neq n$ or if σ does not move any set of size n. Hence $\{A_n | n \in \omega\} \in HWS$ since for any $\sigma \in G$ and any $i \in \omega$ larger than the size of the largest set moved by σ, $j^i\sigma A_n = A_n$ for all n. Furthermore if σ is a transposition swapping a set of size n and a set of size $n+1$, then $j^n\sigma A_n \neq A_n$. Thus $\{A_n | n \in \omega\} \notin HSS$.

1.3 A copy of V inside HSS

If Choice holds in the enveloping model, it is now easy to show that there is an essentially isomorphic structure inside the hereditarily symmetric sets, where the isomorphism does not exists formally inside V but can still be seen from outside. From now on we will restrict our discussion to HSS, since the proofs and results for HWS are exactly the same - we rely on earlier results, which were proved for both classes, rather than the internal structure of each class.

Remark 16. *Until the end of this section, we assume that $G \cap HSS \in \mathcal{F}$, and that the Axiom of Choice holds in V.*

Definition 10. *A BFEXT is a well-founded extensional relation with a top element, namely an element t such that any other element is at the end of a finite descending chain from t.*

Remark 17. *The definitions of top element and BFEXT are preserved under isomorphisms of relations, as long as the graph of the isomorphism is a set. The top element is unique if it exists in a well-founded relation.*

Remark 18. *The definition of a BFEXT can be carried out in any model of strZF as a stratified formula:*

To say that t is the top element of R, we require that $\operatorname{Dom} R$ is the smallest set containing t and closed under R^{-1}. In other words any object in $\operatorname{Dom} R$ must be accessible via a downward path from the top element, and this ensures the top element is unique by well-foundedness.

Since the graph of R is a set, the formula xRy is stratified and homogeneous (i.e. x and y have the same type). This means both the formula asserting that R is well-founded and the above definition of the top element are stratified. Hence the definition of a BFEXT is stratified.

Remark 19. *This is not the same as Roland Hinnion's definition of BFEXTs in [Hinnion 1]. Hinnion does not require the whole domain to be accessible from the top element, thus a BFEXT in his sense may have more than one top element. However our definition does coincide with what Hinnion calls Ω, which is a subclass of his BFEXTs.*

Let $\mathcal{B}$ be the class of all BFEXTs in HSS.

Definition 11. *If $R \in \mathcal{B}$ and $y \in \operatorname{Dom} R$, let R_y be the restriction of R to y and everything below (i.e. the carrier set of R_y is the smallest set containing y and closed under R^{-1}).*

Define an interpretation for the language of set theory on $\mathcal{B}$ as follows - the idea is to look at each BFEXT as a graph representation of some set, with the top element standing in for the set in question:

- $=$ is taken to be isomorphism between BFEXTs
- $\in$ is taken to be $\lhd$, where $R \lhd S$ if and only if R is isomorphic to S_y for some y directly below the top element of S (i.e. $\langle y, s\rangle \in S$ where s is the top element of S).

Any $R \in \mathcal{B}$ is a well-founded extensional relation in V, and so has a Mostowski collapse $m(R)$. Then $m(R) = TC\{x\}$, where x is the image of the top element of R. By uniqueness, we can call $x = t(R)$.

Lemma 20. *Any $x \in V$ is $t(R)$ for some $R \in \mathcal{B}$.*

Proof. This is where we need Choice in V. Let κ be an ordinal the same size as $TC\{x\}$ and λ be a limit ordinal whose cofinality is greater than rank G. Let $X := \{HSS \cap V_{\lambda+\alpha} : \alpha < \kappa\}$. Then $X \in HSS$ and is uniformly symmetric (see Lemma 12 and Corollary 3), so we can copy the membership relation on $TC\{x\}$ over to a relation $R \in HSS$ on the carrier set X. $\square$

Lemma 21. *Let $R, S \in \mathcal{B}$. Then $R \cong S$ if and only if $t(R) = t(S)$, and $R \lhd S$ if and only if $t(R) \in t(S)$.*

Proof. $R \cong S$ if and only if they have the same Mostowski collapse if and only if $t(R) = t(S)$.

If $R \lhd S$: Let s be the top element of S, rSs and $R \cong S_r$. Let y be the Mostowski collapse of S and z be the image of r under the Mostowski map. Then $z \in t(S)$ and $z = t(S_r)$. But $R \cong S_r$ so by the first part $z = t(R)$ too.

If $t(R) \in t(S)$, let r be the preimage of $t(R)$ under the Mostowski map for S. Then rSs, and $t(S_r)$ is the image of r under the Mostowski map, which is just $t(R)$. But then $R \cong S_r$ by the first part, so $R \lhd S$. $\square$

It is now easy to see that $R \leftrightarrow t(R)$ defines an informal isomorphism between $(\mathcal{B}, \cong, \lhd)$ and $(V, =, \in)$. Though we cannot describe this isomorphism formally within V, at least this gives us an easy proof of elementary equivalence between V and $\mathcal{B}$:

Lemma 22. *Let $\phi(\vec{x})$ be a formula in the language of set theory and $\vec{R}$ a tuple in $\mathcal{B}$. Then $\phi(\vec{R})$ holds in $\mathcal{B}$ if and only if $\phi(t(\vec{R}))$ holds in V.*

Proof. We prove the result by induction on the complexity of ϕ.

The atomic cases are true by the previous lemma, and the induction is trivial for conjunction and negation.

Now suppose $\phi \Leftrightarrow_{df} (\exists y)\psi(\vec{x}, y)$.

If $\phi(\vec{R})$ holds in $\mathcal{B}$, there exists $S \in \mathcal{B}$ such that $\psi(\vec{R}, S)$ holds in $\mathcal{B}$. By the induction hypothesis $\psi(t(\vec{R}), t(S))$ holds in V, so $\phi(t(\vec{R}))$ holds in V.

Conversely if $\phi(t(\vec{R}))$ holds in V, then $\psi(t(\vec{R}), y)$ holds in V for some $y \in V$. Let $y = t(S)$ for some $S \in \mathcal{B}$ by Lemma 20, then by the induction hypothesis $\psi(\vec{R}, S)$ holds in $\mathcal{B}$, and so does $\phi(\vec{R})$. $\square$

Corollary 7. *$\mathcal{B}$ satisfies the same first-order sentences as V, and in particular is a model of ZFC (in the sense that it satisfies the new interpretation of all ZFC axioms).*

Remark 20. *If we are willing to look at V and $\mathcal{B}$ as sets (i.e. from an external point of view), we can formalise an isomorphism between V and the quotient $\mathcal{B}/\cong$. Then the above results can be extended to infinitary languages.*

1.3.1 Without the Axiom of Choice

If Choice does not hold in V, there is no guarantee that we can replicate the membership graph of any set inside the BFEXT structure. However we can still show that the BFEXT structure is a model of ZF. The discussion hinges on variants of the following axiom.

Axiom 1. *(IO) Every set is the same size as a set of singletons.*

There is an unpublished result by Thomas Forster that the class of BFEXTs in a model of strZF and IO interprets ZF. Nathan Bowler has proved that IO holds for Forster's original class HS; and a trivial adaptation of his proof generalises the result to HSS if the strong closure condition holds, each permutation in G is finite, and $\mathcal{F}$ is a principal filter. However the proof requires Choice to hold in V, and without it the status of IO in HSS, or even in HS, is not yet clear. Nevertheless, if $G \cap HSS \in \mathcal{F}$, it turns out that just enough of IO can be salvaged to carry out the interpretation.

We will take a detour to give a more general proof, which will give us not just an interpretation of ZF with the BFEXTs, but also allow us to model anti-foundation with the class of accessible pointed graphs.

2 Interpreting ZF in stratified theories

2.1 The general construction

Definition 12. *Let strZF$^-$ be the weakening of strZF where Foundation is replaced by the axiom that there is no universal set, i.e.*

$$(\forall x)(\exists y) y \notin x$$

In this section we work in a model M of strZF$^-$, plus a few other axioms to be specified.

First some stratified definitions for handling relations:

Definition 13. *Let R be a relation, $x \in \operatorname{Dom} R$ and $X \subset \operatorname{Dom} R$:*

- $R^{-1} := \{\langle y, x\rangle : \langle x, y\rangle \in R\}$.
- $R \restriction X := R \cap X^2$.
- $Trans(R) := \bigcap\{S \subset (\operatorname{Dom} R)^2 : R \subset S \wedge (\forall x, y, z \in \operatorname{Dom} S)(xSy \wedge ySz \Rightarrow xSz)\}$, *the transitive closure of R.*
- $x_R := \{x\} \cup \{y \in \operatorname{Dom} R : \langle y, x\rangle \in Trans(R)\}$, *the closure of $\{x\}$ under R^{-1}.*
- $R_x := R \restriction x_R$.

Definition 14. *An* accessible pointed graph *(APG) is a directed graph with a distinguished node called its* point*, such that there is a directed path from the point to any other node.*

Remark 21. *It is obvious that every BFEXT is an APG if we regard the unique top element as the distinguished point. Furthermore for any BFEXT one can easily recover the distinguished point from the relation itself. However this is not true for non-well-founded APGs, and it is easy to find examples of directed graphs accessible from more than one point. Thus for a general APG we will explicitly specify its distinguished point in addition to the graph relation.*

Formally, let A be an APG if $A = \langle R, \{\{\{r\}\}\}\rangle$ such that:

- R is a relation.
- $r \in \operatorname{Dom} R$.
- $\operatorname{Dom} R = r_R$.

Here R represents the graph and r the point - the triple singleton is to make sure the predicate "A is an APG" is stratified. $\langle x, y\rangle \in R$ denotes a directed edge from y to x and informally represents "$x \in y$". For convenience we write $[R, r]$ for such an object.

Let $\mathcal{G}$ be a definable subclass of APGs with a stratified definition. We call members of $\mathcal{G}$ *standard pointed graphs* (SPGs). The intention is to use these graphs to represent the membership graphs of sets in a ZF model. Further properties of the resulting model will depend on the class $\mathcal{G}$.

Added to our theory is a weakened version of IO (Axiom 1):

Axiom 2. *(Weak IO for $\mathcal{G}$) If $[R, r] \in \mathcal{G}$, then $\operatorname{Dom} R$ is the same size as a set of singletons.*

Definition 15. *If $\mathcal{G}$ is a definable class of APGs, let the theory strZF$_{\mathcal{G}}$ be strZF^{-} plus Weak IO for $\mathcal{G}$.*

Definition 16. *Let $[R, r]$ be an APG and $\mathbf{n}$ a concrete natural number such that $[R_x, x] \in \mathcal{G}$ whenever there is a chain $x = x_{\mathbf{0}}\ R \ldots R\ x_{\mathbf{n}} = r$, i.e. there is a directed path of length $\mathbf{n}$ from the distinguished point r to x. Then we call $[R, r]$ an* extended pointed graph *(EPG).*

Remark 22. *Intuitively, each EPG is built up by grouping disjoint SPGs together in a tree of finite depth whose branches all have equal lengths. At the moment we specify $\mathbf{n}$ as a concrete natural number since there is no need quantify over $\mathbf{n}$; in fact $n \leq 2$ will be enough for our purposes.*

We now state the extra axioms necessary to interpret ZF in $\mathcal{G}$.

Axiom 3. *(Axiom of Preservation for $\mathcal{G}$) If $[R, r] \in \mathcal{G}$, $[S, s]$ is an APG and $\phi(x, y)$ is a stratified formula (possibly with parameters) such that*

$$\begin{aligned}
&(\forall x \in \operatorname{Dom} R)(\exists! y \in \operatorname{Dom} S)\phi(x, y) \wedge \\
&(\forall y \in \operatorname{Dom} S)(\exists! x \in \operatorname{Dom} R)\phi(x, y) \wedge \\
&(\forall x, y \in \operatorname{Dom} R)(\forall z, t \in \operatorname{Dom} S)\Big((\phi(x, z) \wedge \phi(y, t)) \Rightarrow (xRy \Leftrightarrow zSt)\Big)
\end{aligned}$$

Then $[S, s] \in \mathcal{G}$.

Axiom 4. *(Axiom of Stability for $\mathcal{G}$) If $[R, r] \in \mathcal{G}$ and f is an isomorphism between R_x and R_y for some $x, y \in \operatorname{Dom} R$, then f is the identity.*

Axiom 5. *(Axiom of Quotient for $\mathcal{G}$) Let $[R, r]$ be an EPG. Then there exists an SPG $[Q, q]$ and a surjective quotient map $\pi : \operatorname{Dom} R \to \operatorname{Dom} Q$ such that*

$$(\forall z, t \in \operatorname{Dom} Q)(zQt \Leftrightarrow (\exists x, y \in \operatorname{Dom} R)(z = \pi(x) \wedge t = \pi(y) \wedge xRy))$$

and $\pi(r) = q$.

Furthermore, if $x \in \operatorname{Dom} R$ *and* $[R_x, x]$ *is an SPG, then* π *is an isomorphism between* R_x *and* $Q_{\pi(x)}$.

Remark 23. *The Axiom of Preservation essentially states that* $\mathcal{G}$ *is closed under external isomorphisms; but if* x *and* y *have different types in* ϕ*, then we cannot show that the graph of* ϕ *is a set with only stratified axioms. The Axiom of Stability will ensure that our structure for set theory is extensional, whereas the Axiom of Quotient will be useful time and again in building new SPGs.*

Now let M be a model of strZF$^-$ and $\mathcal{G}$ be a stratified definable class provable in strZF$^-$ to be a class of APGs. Assume that M also satisfies the Axioms of Weak IO, Preservation, Stability and Quotient for $\mathcal{G}$.

Let $\mathcal{G}_M$ be the class $\mathcal{G}$ of SPGs in the model M. Reinterpret the language of set theory on $\mathcal{G}_M$ as follows:

- The equality relation is $[R, r] \equiv [S, s]$ if and only if there is an isomorphism $R \leftrightarrow S$ which sends r to s.
- The membership relation is $[R, r] \lhd [S, s] \Leftrightarrow_{df} (\exists x S s)[R, r] \equiv [S_x, x]$.

It is easy to see that $\equiv$ is an equivalence relation respected by $\lhd$.

If ϕ is a formula in the language of set theory (possibly unstratified), let $\bar{\phi}$ be the formula obtained from ϕ by replacing all $=$ with $\equiv$, $\in$ with $\lhd$ and $\exists x$ with $\exists x \in \mathcal{G}_M$. Then $\bar{\phi}$ is stratified since both $\equiv$ and $\lhd$ have stratified definitions, and it is the interpretation of ϕ in the structure $\mathcal{G}_M$. We write $\mathcal{G}_M \models \phi$ to mean $M \models \bar{\phi}$.

We now prove all axioms of ZF except Foundation and Infinity in this interpretation.

Lemma 23. *Let* $[R, r] \in \mathcal{G}_M$ *and* $\mathbf{n}$ *be a concrete natural number. Then* $\operatorname{Dom} R$ *has the same size as* $\{\iota^{\mathbf{n}} x : x \in \operatorname{Dom} S\}$ *for some* $[S, s] \in \mathcal{G}_M$.

Proof. By induction on $\mathbf{n}$.

Suppose the result is true for $\mathbf{n}$. Let $\pi : \{\{x\} : x \in X\} \leftrightarrow \operatorname{Dom} R$ be a bijection for some $X \in M$.

Define a relation P on X by $xPy \Leftrightarrow_{df} \pi(\{x\}) R \pi(\{y\})$.

Then $[P, \bigcup \pi^{-1} r] \in \mathcal{G}_M$ by the Axiom of Preservation, so X has the same size as $\{\iota^{\mathbf{n}} x : x \in \operatorname{Dom} S\}$ for some $[S, s] \in \mathcal{G}_M$. Therefore $\operatorname{Dom} R$ has the same size as $\{\iota^{\mathbf{n+1}} x : x \in \operatorname{Dom} S\}$. $\square$

Lemma 24. *Extensionality holds.*

Proof. Suppose $[R, r]$ and $[S, s]$ have the same $\lhd$-predecessors. The set F of all isomorphisms $R_x \leftrightarrow S_y$ where xRr, ySs is given by stratified comprehension from $\mathcal{P}(\operatorname{Dom} R \times \operatorname{Dom} S)$.

If $f, g \in F$ are defined on the same $x \in \operatorname{Dom} R$, then they are both defined on x_R and $\{\langle f(y), g(y)\rangle : y \in x_R\}$ is an isomorphism between $S_{f(x)}$ and $S_{g(x)}$; thus by the Axiom of Stability it is the identity. In particular $f(x) = g(x)$.

Hence $\bigcup F$ is a function, and by the same argument in the other direction $\bigcup F$ is injective. By the hypothesis

$$\operatorname{dom} \bigcup F = \operatorname{Dom} R \setminus \{r\} \wedge \operatorname{ran} \bigcup F = \operatorname{Dom} S \setminus \{s\}$$

so $\bigcup F \cup \{\langle r, s\rangle\}$ is an isomorphism $R \leftrightarrow S$ that sends r to s. $\square$

The next result gives us much freedom in creating new SPGs:

Lemma 25. *(Supertransitivity Lemma) Let $A \in M$ be a set of SPGs. Then there is an SPG $[S, s]$ such that*

$$(\forall [P, p] \in \mathcal{G}_M)\Big([P, p] \lhd [S, s] \Leftrightarrow (\exists [Q, q] \in A)[P, p] \equiv [Q, q]\Big)$$

Proof. For any SPG $[P, p]$ and $x \in \operatorname{Dom} P$, let $x^P := \langle \iota^3 x, P\rangle$ and

$$P^P := \{\langle x^P, y^P\rangle : \langle x, y\rangle \in P\}$$

P^P exists by stratified Comprehension from $\mathcal{P}^4(\{P\} \cup \operatorname{Dom} P)$. Then $[P^P, p^P]$ is an SPG by the Axiom of Preservation, since the map $x \mapsto x^P$ is stratified.

By stratified Comprehension let

$$A^* := \{[P^P, p^P] : [P, p] \in A\}$$

By construction the $\operatorname{Dom} P^P$ are disjoint for all $[P^P, p^P] \in A^*$, since each x^P is an ordered pair with second component P. Note that the formula $X = A^*$ is stratified, and going from x to x^P or A to A^* raises type by 5.

Let $b \notin \bigcup\{\operatorname{Dom} P : [P, p] \in A^*\}$ and

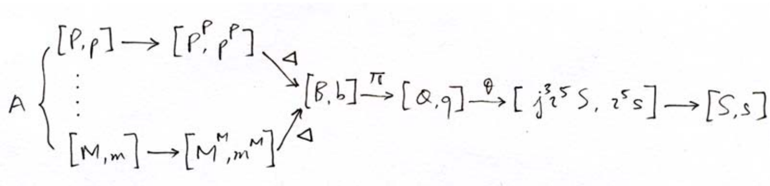
$$B := \bigcup\{P : [P, p] \in A^*\} \cup (\{p : [P, p] \in A^*\} \times \{b\})$$

Then $[B, b]$ is an EPG, so by the Axiom of Quotient let π be the quotient map from $[B, b]$ to an SPG $[Q, q]$.

By Lemma 23, there is a bijection $\theta : \operatorname{Dom} Q \leftrightarrow \iota^5 “H$ for some $H \in M$. Define a relation on H by

$$S := \{\langle x, y\rangle : \langle \theta^{-1}\iota^5 x, \theta^{-1}\iota^5 y\rangle \in Q\}$$

and let $\theta(q) = \iota^5 s$. Then $[S, s]$ is an SPG by the Axiom of Preservation. Below is an informal diagram of the construction:



Let $[P, p] \in A$. Since $[P^P, p^P] = [B_{p^P}, p^P]$ is an SPG, by the Axiom of Quotient the restriction of π induces an isomorphism $[P^P, p^P] \leftrightarrow [Q_{\pi(p^P)}, \pi(p^P)]$.

By stratified Comprehension define

$$\sigma := \{\langle x, y\rangle \in \operatorname{Dom} P \times \operatorname{Dom} S : \theta\pi(x^P) = \iota^5 y\}$$

Then σ is an isomorphism between P and $S_\sigma(p)$, therefore

$$[P, p] \equiv [S_\sigma(p), \sigma(p)] \lhd [S, s]$$

Conversely let $[R, r] \lhd [S, s]$; without loss of generality let $[R, r] = [S_y, y]$ where ySs.

Then there exists xBb such that $\theta\pi(x) = \iota^5 y$, but then $x = p^P$ for some $[P, p] \in A$. Thus by the argument above, $[P, p] \equiv [S_y, y]$. ☐

Lemma 26. *Comprehension holds.*

Proof. Let $[R, r]$ be an SPG and $\phi(x)$ a formula in the language of set theory with parameters in $\mathcal{G}_M$, then $\bar{\phi}$ is stratified.

Hence by stratified Comprehension from $\mathcal{P}(R) \times \iota^3 “ \operatorname{Dom} R$, the following set exists in M

$$A := \{[R_x, x] : xRr \wedge \bar{\phi}[R_x, x]\}$$

By the Supertransitivity Lemma, there exists $[S, s]$ such that for any SPG $[P, p]$

$$[P, p] \lhd [S, s] \Leftrightarrow (\exists [Q, q] \in A)([P, p] \equiv [Q, q]) \Leftrightarrow ([P, p] \lhd [R, r] \wedge \bar{\phi}[P, p])$$

☐

Lemma 27. *Pairing holds.*

Proof. Given $[R, r]$ and $[S, s]$, let $A := \{[R, r], [S, s]\}$ and the result holds by the Supertransitivity Lemma. $\square$

Lemma 28. *Union holds.*

Proof. Given $[R, r]$, by stratified Comprehension from $\mathcal{P}R$ and $\mathrm{Dom}\, R$ we get

$$A := \{[R_x, x] : (\exists y \in \mathrm{Dom}\, R)(xRy \wedge yRr)\}$$

By the Supertransitivity Lemma, there exists $[S, s]$ such that

$$\begin{aligned}[P, p] \lhd [S, s] &\Leftrightarrow (\exists [Q, q] \in A)[P, p] \equiv [Q, q] \\ &\Leftrightarrow (\exists [T, t] \lhd [R, r])[P, p] \lhd [T, t]\end{aligned}$$

$\square$

Lemma 29. *Power Set holds.*

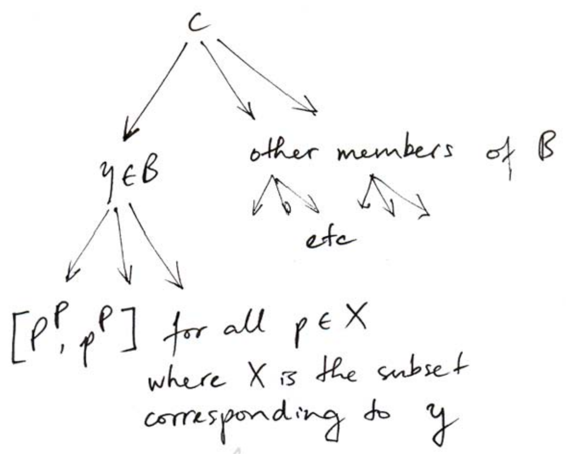
Proof. Given $[R, r]$, let $A := \{[R_x, x] : xRr\}$ and construct A^* like in the proof of the Supertransitivity Lemma.

Assume there exist B disjoint from $\bigcup\{\mathrm{Dom}\, P : [P, p] \in A^*\}$ and a bijection ϕ from B to ι^4"$\mathcal{P}\{x \in \mathrm{Dom}\, R : xRr\}$. We form an EPG representing the power set of $[R, r]$ by letting each member of B stand for the corresponding subset of $[R, r]$.

Let $c \notin \bigcup\{\mathrm{Dom}\, P^P : [P^P, p^P] \in A^*\}$ be a new vertex, and define a stratified relation:

$$C := \bigcup\{P^P : [P^P, p^P] \in A^*\} \cup \left\{\langle p^P, y\rangle : [P^P, p^P] \in A^* \wedge p \in \bigcup^4 \phi(y)\right\} \cup (B \times \{c\})$$

Below is an illustration of the graph C.



Then $[C, c]$ is an EPG, and by the Axiom of Quotient there is a quotient map π from $[S, s]$ to an SPG $[Q, q]$.

Since $[Q, q]$ is an SPG, by Lemma 23 there is a bijection $\theta : \operatorname{Dom} Q \leftrightarrow \iota^5 ``G$ for some $G \in M$. Define a relation

$$S := \{\langle x, y\rangle : \langle \theta^{-1}\iota^5 x, \theta^{-1}\iota^5 y\rangle \in Q\}$$

and let $\theta(q) = \iota^5 s$. Then $[S, s]$ is an SPG by the Axiom of Preservation.

We now prove that

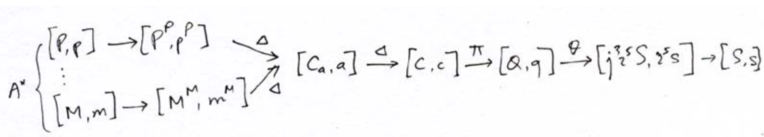
$$(\forall [P, p] \in \mathcal{G}_M)([P, p] \lhd [S, s] \Leftrightarrow \mathcal{G}_M \models [P, p] \subset [R, r])$$

Suppose $\mathcal{G}_M \models [P, p] \subset [R, r]$. Let

$$a := \phi^{-1}\iota^4 ``\{xRr : [R_x, x] \lhd [P, p]\}$$

and let b be such that $\theta\pi(a) = \iota^5 b$.

We have the following informal diagram:



If $[T, t] \lhd [P, p]$, then $[T, t] \equiv [R_Z, z]$ for some $z \in \bigcup^4 \phi(a)$.

For convenience write $Z := R_z$, then $[Z^Z, z^Z]$ is an SPG so

$$[Z^Z, z^Z] \equiv [Q_{\pi(z^Z)}, \pi(z^Z)] \lhd [Q_{\pi(a)}, \pi(a)]$$

Define by stratified Comprehension an isomorphism between Z and $S_\sigma(z)$ as follows

$$\sigma := \{\langle x, y\rangle \in \operatorname{dom} Z \times \operatorname{Dom} S : \theta\pi(x^Z) = \iota^5 y\}$$

Hence $[Z, z] \equiv [S_\sigma(z), \sigma(z)] \lhd [S_b, b]$.

On the other hand if xSb, then $\iota^5 x = \theta\pi(z^Z)$ where z is such that $\iota^4 ``z \in \phi(a)$ and $Z = R_z$.

Then $[Z, z] \lhd [P, p]$ by definition of a, but $[Z, z] \equiv [S_x, x]$ by the same argument as in the last paragraph, so $[S_x, x] \lhd [P, p]$.

By Extensionality we have proved

$$[P, p] \equiv [S_b, b] \lhd [S, s]$$

Conversely let $[P, p] \lhd [S, s]$, then $[P, p] \equiv [S_b, b]$ for some bSs.

Let xSb, then $\iota^5 x = \theta\pi(z^Z)$ where zRr and $Z = R_z$. By the same reasoning as above we have $[S_x, x] \equiv [Z, z] \lhd [R, r]$. Therefore

$$\mathcal{G}_M \models [P, p] \subset [R, r]$$

and thus $[S, s]$ is the power set of $[R, r]$ in $\mathcal{G}_M$.

To complete the proof we describe how to construct B as required. First let

$$C := \mathcal{P}\{x \in \operatorname{Dom} R : xRr\}$$

Let $c \notin \bigcup^4\{\operatorname{Dom} P : [P, p] \in A^*\} \cup \operatorname{Dom} R$ and define

$$D := \{\{\{\langle x, c\rangle : x \in y\}\} : y \in C\}$$

Then D has a natural stratified bijection with $\iota^4 \text{“} C$. Now let

$$B := D \setminus \emptyset \cup \{\iota^4 c\}$$

Then B has the same size as D since $\iota^4 c \neq \{\{\langle x, c\rangle\}\}$ for any $x \in \operatorname{Dom} R$, and B is disjoint from $\bigcup\{\operatorname{Dom} P : [P, p] \in A^*\}$ since $(\forall x \in B) c \in \bigcup^4 x$. □

Lemma 30. *Collection holds.*

Proof. Let $[R, r] \in \mathcal{G}_M$ and $\phi(x, y)$ be a formula in the language of set theory such that

$$\mathcal{G}_M \models (\forall x \in [R, r])(\exists y)\phi(x, y)$$

For any $[P, p] \lhd [R, r]$, there exists $x \in \operatorname{Dom} R$ such that $[P, p] \equiv [R_x, x]$. Thus by stratified Collection in M we have a set A such that

$$(\forall [P, p] \lhd [R, r])(\exists [Q, q] \in A)([Q, q] \in \mathcal{G}_M \wedge \bar{\phi}([P, p], [Q, q])$$

By stratified Comprehension let $B := A \cap \mathcal{G}_M$. By the Supertransitivity Lemma, let $[S, s]$ be such that for all SPG $[P, p]$

$$[P, p] \lhd [S, s] \Leftrightarrow (\exists [Q, q] \in B)[P, p] \equiv [Q, q]$$

So for any $[P, p] \lhd [R, r]$ there exists $[Q, q] \lhd [S, s]$ such that $\bar{\phi}([P, p], [Q, q])$ holds. □

Remark 24. *Stratified Collection in M is only used in the proof of Collection in $\mathcal{G}_M$. Therefore we could drop stratified Collection in M if we do not need Collection in the new interpretation.*

2.2 The BFEXTs and Foundation

It is clear that any BFEXT is an APG if we regard the top element as the distinguished point, and recall from Remark 18 that the class of BFEXTs has a stratified definition.

Let M be a model of strZF_{BFEXT} (i.e. strZF^- plus Weak IO for BFEXTs) and write $\mathcal{B}_M$ for the class of BFEXTs in M. We will prove the following result:

Theorem 2. *If M is a model of $strZF_{BFEXT}$, then the class $\mathcal{B}_M$ of BFEXTs in M models ZF in the given interpretation.*

Lemma 31. *(Axiom of Preservation for BFEXTs) If $[R, r] \in \mathcal{B}_M$, $[S, s]$ is an APG and $\phi(x, y)$ is a stratified formula (possibly with parameters) such that*

$$(\forall x \in \operatorname{Dom} R)(\exists! y \in \operatorname{Dom} S)\phi(x, y) \wedge (\forall y \in \operatorname{Dom} S)(\exists! x \in \operatorname{Dom} R)\phi(x, y) \wedge$$
$$(\forall x, y \in \operatorname{Dom} R)(\forall z, t \in \operatorname{Dom} S)\Big((\phi(x, z) \wedge \phi(y, t)) \Rightarrow (xRy \Leftrightarrow zSt)\Big)$$

Then $[S, s] \in \mathcal{B}_M$.

Proof. Let $\emptyset \neq A \subset \operatorname{Dom} S$, then

$$B := \{x \in \operatorname{Dom} R : (\exists y \in A)\phi(x, y)\} \neq \emptyset$$

Let b be R-minimal in B, then there exists $a \in A$ such that $\phi(b, a)$ holds. Then a is S-minimal in A, so S is well-founded.

Let $a, b \in \operatorname{Dom} S$ such that $\{x : xSa\} = \{x : xSb\}$. Let $c, d \in \operatorname{Dom} R$ such that $\phi(c, a)$, $\phi(d, b)$ hold. Then $\{x : xRc\} = \{x : xRd\}$ so $c = d$ and $a = b$, i.e. S is extensional. □

Lemma 32. *(Axiom of Stability for BFEXTs) If $[R, r] \in \mathcal{B}_M$ and f is an isomorphism between R_x and R_y for some $x, y \in \operatorname{Dom} R$, then f is the identity.*

Proof. If $\{z \in \operatorname{Dom} R : f(z) \neq z\} \neq \emptyset$, let a be its R-minimal element. Then $(\forall zRa) f(z) = z$, but f is bijective so $\{z : zRf(a)\} = \{f(z) : z \in a\} = \{z : zRa\}$. Hence $a = f(a)$, contradicting the choice of a. □

Lemma 33. *(Axiom of Quotient for BFEXTs) Let $[R, r]$ be a well-founded APG. Then there exists a BFEXT $[Q, q]$ and a surjective quotient map $\pi : \operatorname{Dom} R \to \operatorname{Dom} Q$ such that*

$$(\forall z, t \in \operatorname{Dom} Q)\Big(zQt \Leftrightarrow (\exists x, y \in \operatorname{Dom} R)(z = \pi(x) \wedge t = \pi(y) \wedge xRy)\Big)$$

and $\pi(r) = q$.

Furthermore, if $x \in \operatorname{Dom} R$ and $[R_x, x]$ is a BFEXT, then π is an isomorphism between R_x and $Q_{\pi(x)}$.

Proof. Define a relation $\sim$ on $\operatorname{Dom} R$ by recursion on R as follows:

If T is a relation, let $\phi(T)$ be the formula

$$(\forall x \in \operatorname{Dom} T)(\forall y \in \operatorname{ran} T)(xTy \Leftrightarrow ((\forall zRx)(\exists tRy)zTt \wedge (\forall tRy)(\exists zRx)zTt))$$

Now define

$$x \sim y \Leftrightarrow_{df} (\exists T \subset X^2)(T \text{ is a relation} \wedge \phi(T) \wedge xTy)$$

We will show that

$$(\forall x, y \in \operatorname{Dom} R)\Big(x \sim y \Leftrightarrow \Big((\forall zRx)(\exists tRy)zTt \wedge (\forall tRy)(\exists zRx)zTt\Big)\Big) \qquad (\dagger)$$

First suppose $x \sim y$, then there is a relation $T \subset \operatorname{Dom} R^2$ such that xTy and $\phi(T)$ holds.

If zRx, then by $\phi(T)$ there is some tRy such that zTt, but then $z \sim t$ by definition. Similarly if tRy then there is zRx such that zTt and so $z \sim t$. Thus the left to right implication holds.

Conversely let $\bar{R} := Trans(R)$, then $\bar{R}$ is well-founded. Suppose there is some $x, y \in \operatorname{Dom} R$ such that

$$(\forall zRx)(\exists tRy)(s \sim t) \wedge (\forall tRy)(\exists zRx)(s \sim t) \wedge (x \not\sim y)$$

The set of all such x, y exists by Union and stratified Comprehension from $\mathcal{P}^2 \operatorname{Dom} R$. Let a be the $\bar{R}$-minimal such x and b be $\bar{R}$-minimal value of y corresponding to a. Define

$$T := \{\langle x, y\rangle : x \sim y \wedge x\bar{R}a \wedge y\bar{R}b\} \cup \{\langle a, b\rangle\}$$

Let $x \in \operatorname{dom} T$ and $y \in \operatorname{ran} T$. We show that

$$xTy \Leftrightarrow \Big((\forall zRx)(\exists tRy)zTt \wedge (\forall tRy)(\exists zRx)zTt\Big)$$

The left to right implication holds by the previous paragraph, since the restriction of T to $\{x : x\bar{R}a\} \times \{y : y\bar{R}b\}$ is the same as $\sim$.

The right to left implication holds if $x\bar{R}a$ by minimality of a, if $x = a$ and $y\bar{R}b$ by minimality of g, and if $x = a$, $y = b$ by default.

Thus $\phi(T)$ holds, and we have $x \sim y$ contradicting our earlier assumption. We have proved the other direction of $(\dagger)$.

Now we show that $\sim$ is an equivalence relation.

Suppose x is R-minimal such that $x \not\sim x$, then $(\forall y R x)(y \sim y)$, so $x \sim x$ by (†).

Also it is clear that $x \sim y$ if and only if $y \sim x$; for if T witnesses $x \sim y$, then T^{-1} witnesses $y \sim x$.

Suppose $x \sim y \sim z$ but $x \not\sim z$. Choose y to be R-minimal by stratified Comprehension. If sRx, then $s \sim v \sim t$ for some vRy, tRz but y is R-minimal so $s \sim t$. Similarly if tRz then $s \sim t$ for some sRx, so $x \sim z$ i.e. contradiction. Therefore $\sim$ must be transitive.

Now we can use stratified Comprehension on $\mathcal{P}\,\mathrm{Dom}\,R$ to get the set A of $\sim$-equivalence classes.

For any $x \in \mathrm{Dom}\,R$, write $[x] \in A$ for the $\sim$-equivalence class of x. Define a relation Q on A by

$$yQt \Leftrightarrow_{df} (\exists x \in y)(\exists z \in t)xRz$$

Let $B \subset A$ and let x be R-minimal in $\bigcup B$, then $[x]$ is Q-minimal in B. So Q is well-founded.

Let $a, b \in A$ such that $\{x : xQa\} = \{x : xQb\}$.

Let $x \in a$ and $y \in b$. If zRx, then $[z]Qa$ so $[z]Qb$. Thus zRs for some $s \sim y$ by definition of Q. But then by (†) $z \sim t$ for some tRy. Similarly if tRy, then $z \sim t$ for some zRx. By the other direction of (†) we have $x \sim y$.

Hence $a = b$, which shows that Q is extensional.

Thus A has a bijection θ with a set B of singletons. Let $Y := \bigcup B$ and let

$$\pi := \{\langle x, y\rangle \in \mathrm{Dom}\,R \times Y : (\exists a \in A)(x \in a \wedge y \in \theta(a))\}$$

Define a relation S on Y by

$$ySz \Leftrightarrow_{df} \{x : x\theta y\}Q\{x : x\theta z\}$$

Then S is also well-founded and extensional. It is clear that π is a surjection, $xRy \Rightarrow \pi(x)S\pi(y)$, and

$$pSq \Rightarrow (\exists x, y \in \mathrm{Dom}\,R)(p = \pi(x) \wedge q = \pi(y) \wedge xRy)$$

Suppose $C \subset Y$ contains $\pi(r)$ and is closed under S^{-1}, then $\{x \in \mathrm{Dom}\,R : \pi(x) \in C\}$ contains p and is closed under R^{-1}. Thus $C = Y$, so $\pi(r)$ is the top element of Y.

Now let $[R_x, x]$ be a BFEXT for some $x \in \mathrm{Dom}\,R$. We need to show that π is injective on x_R.

Suppose there are $y \neq z \in x_R$ such that $y \sim z$, and choose y to be R-minimal. For any $t \in x_R$ with tRy, $t \sim p$ for some pRz. Then $p \in x_R$ and y is R-minimal so $t = p$. Similarly tRy if tRz, so by extensionality of R on Z we have $y = z$. Contradiction. $\square$

Remark 25. *What we just proved is slightly stronger than the Axiom of Quotient as previously stated. If we take SPGs to be BFEXTs, then the corresponding EPGs are well-founded APGs, but not all well-founded APGs are EPGs.*

In light of the results in section 2.1, $\mathcal{B}_M$ is thus a model of ZF minus Foundation and Infinity.

Lemma 34. *Infinity holds for $\mathcal{B}_M$.*

Proof. Let $R \in M$ be a non-empty well-order with no maximal element, and let $s \notin \operatorname{Dom} R$. Define $S := R \cup (\operatorname{Dom} R \times \{s\})$, then $[S, s]$ is a BFEXT. But the set $\{[S_x, x] : x \in \operatorname{Dom} R\}$ is well-ordered by $\lhd$, so $[S, s]$ is a nonzero limit von Neumann ordinal in $\mathcal{B}_M$. $\square$

Lemma 35. *Foundation holds for $\mathcal{B}_M$.*

Proof. Let $[R, r]$ be a BFEXT, and let x be R-minimal in $\{x : xRr\}$. Suppose $[P, p] \lhd [R, r]$ and $[P, p] \lhd [R_x, x]$, then there is some yRr such that $[R_y, y] \equiv [P, p] \lhd [R_x, x]$. Contradiction. Thus $[R_x, x]$ is a minimal member of R, r in the sense of $\mathcal{B}_M$. $\square$

Thus we have proved Theorem 2

2.3 APGs and anti-foundation

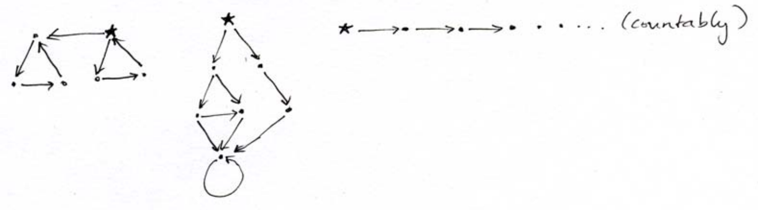
The idea is to consider the class of accessible pointed graphs, in order to get a non-well-founded model of set theory inside HS. We will model the anti-foundation axiom AFA as in [Aczel 1], which states that every APG has a unique decoration, where:

Definition 17. *A* decoration *on an APG $[A, a]$ is a function f defined on* $\operatorname{Dom} A$ *such that* $(\forall x \in \operatorname{Dom} A) f(x) = \{f(y) : y \in x\}$.

Axiom 6. *(AFA) Every APG has a unique decoration.*

Definition 18. *The theory ZFA is ZF with AFA in place of Foundation.*

There are some problems with considering the class of APGs as is. Firstly if HSS does not satisfy IO, then we can easily find an APG $[A, a]$ where $\operatorname{Dom} A$ is not the same size as any set of singletons : If $X \in HSS$ is a counterexample to IO, let $a \notin X$ and set $A := X \times \{a\}$. But more importantly, since AFA states that each APG has a unique decoration, the representation of sets as APGs is not unique: the same set can be represented by infinitely many non-isomorphic APGs. For example these APGs, with the star denoting the distinguished point, all represent the same set according to AFA, namely the unique Quine atom $a = \{a\}$:



Therefore it is advantageous to build some kind of extensionality clause into our class of APG under consideration. This will provide us with canonical pictures of sets and simplify the new identity relation, as well as giving us enough control over these APGs in HSS to prove the necessary version of IO.

We borrow the notion of a bisimulation from computer science:

Definition 19. *Given a relation R, a* bisimulation *on R is a relation $\sim$ on* $\operatorname{Dom} R$ *such that*

$$(\forall x, y \in \operatorname{Dom} R)(x \sim y \Rightarrow ((\forall z R x)(\exists t R y)(z \sim t) \wedge (\forall t R y)(\exists z R x)(z \sim t)))$$

With this, we restrict our attention to the following class of graphs:

Definition 20. *A relation R is* rigid *if any bisimulation on R is the identity. A* rigid pointed graph *(RPG) is an APG $[R, r]$ where R is rigid.*

Remark 26. *If R is a relation and $S \subset R$ is such that*

$$(\forall x \in \operatorname{Dom} S)(\forall y \in \operatorname{Dom} R)(y R x \Rightarrow \langle y, x \rangle \in S)$$

we say S is a closed subset *of R, and then clearly any bisimulation on S is a bisimulation on R.*

Thus closed subsets of rigid relations are rigid. In particular if $[R, r]$ is an RPG, then $[R_x, x]$ is an RPG for any $x \in \operatorname{Dom} R$.

It is clear that the class of RPGs has a stratified definition. Let M be a model of strZF_{RPG} (i.e. strZF^- plus Weak IO for RPGs, as in Definition 15) and let $\mathcal{R}_M$ be the class of RPGs in M.

Lemma 36. *(Axiom of Preservation for RPGs) If $[R, r] \in \mathcal{R}_M$, $[S, s]$ is an APG and $\phi(x, y)$ is a stratified formula (possibly with parameters) such that*

$$(\forall x \in \operatorname{Dom} R)(\exists! y \in \operatorname{Dom} S)\phi(x, y) \wedge (\forall y \in \operatorname{Dom} S)(\exists! x \in \operatorname{Dom} R)\phi(x, y) \wedge$$
$$(\forall x, y \in \operatorname{Dom} R)(\forall z, t \in \operatorname{Dom} S)((\phi(x, z) \wedge \phi(y, t)) \Rightarrow (x R y \Leftrightarrow z S t))$$

Then $[S, s] \in \mathcal{R}_M$.

Proof. Let $\sim$ be a bisimulation on S. Define a relation $\approx$ on $\mathrm{Dom}\, R$ by

$$x \approx y \Leftrightarrow_{df} (\exists z, t \in \mathrm{Dom}\, S)(\phi(x, z) \wedge \phi(y, t) \wedge z \sim t)$$

Then it is easy to see that $\approx$ is a bisimulation on R, therefore it is the identity. Hence by definition of $\approx$, $\sim$ is also the identity. □

Lemma 37. *(Axiom of Stability for RPGs) If* $[R, r] \in \mathcal{R}_M$ *and* f *is an isomorphism between* R_x *and* R_y *for some* $x, y \in \mathrm{Dom}\, R$*, then* f *is the identity.*

Proof. The graph of f regarded as a binary relation itself is a bisimulation on R. Therefore f is the identity. □

Lemma 38. *(Axiom of Quotient for RPGs) Let* $[R, r]$ *be an APG, then there is an RPG* $[S, s]$ *and a surjective map* $\pi : \mathrm{Dom}\, R \to \mathrm{Dom}\, S$ *such that*

$$(\forall z, t \in \mathrm{Dom}\, S)(zSt \Leftrightarrow (\exists x, y \in \mathrm{Dom}\, R)(z = \pi(x) \wedge t = \pi(y) \wedge xRy))$$

and $\pi(r) = s$.

Furthermore, if $p \in \mathrm{Dom}\, R$ *and* $[R_p, p]$ *is an RPG, then* π *is an isomorphism between* R_p *and* $S_{\pi(p)}$.

Proof. We use the construction of strongly extensional quotients (Theorem 2.4 to Lemma 2.17) in [Aczel 1], with simplification when possible and some changes to accommodate our weakened theory:

Define a relation on $\mathrm{Dom}\, R^2$ by stratified Comprehension as

$$x \approx y \Leftrightarrow_{df} (\exists \sim)(\sim \text{ is a bisimulation on } R \wedge x \sim y)$$

If $\sim$ is a relation on $\mathrm{Dom}\, R$, write $\sim^+$ for the relation

$$x \sim^+ y \Leftrightarrow_{df} (\forall z R x)(\exists t R y)(z \sim t) \wedge (\forall t R y)(\exists z R x)(z \sim t)$$

Then $\sim$ is a bisimulation if and only if $\sim \subset \sim^+$. It is clear that if $\sim_1 \subset \sim_2$, then $\sim_1^+ \subset \sim_2^+$, i.e. $\sim$ is monotonic with respect to the subset relation. The following argument is a special case of the Knaster-Tarski theorem [Tarski 1]:

If $x \approx y$, then $x \sim y$ for some bisimulation $\sim$ on R, so $x \sim^+ y$. But $\sim^+ \subset \approx^+$ since $\sim \subset \approx$, so $x \approx^+ y$. Hence $\approx \subset \approx^+$, so $\approx^+ \subset \approx^{++}$. But then $\approx^+$ is a bisimulation, so $\approx^+ \subset \approx$.

We have shown that

$$(\forall x, y \in \operatorname{Dom} R)\Big(x \approx y \Leftrightarrow \Big((\forall zRx)(\exists tRy)(z \approx t) \wedge (\forall tRy)(\exists zRx)(z \approx t)\Big)\Big)$$

Furthermore, $\approx$ is an equivalence relation because:

- The identity on R is a bisimulation, so $\approx$ is reflexive.
- For any bisimulation $\sim$ on R, $(\sim^{-1})^{+} = (\sim^{+})^{-1} \supset \sim^{-1}$, so $\approx$ is symmetric.
- For any bisimulations $\sim_1, \sim_2$ on R, define

  $$x \simeq y \Leftrightarrow_{df} (\exists z \in \operatorname{Dom} R)(x \sim_1 z \wedge z \sim_2 y)$$

  If $x \simeq y$ with z as the witness, then $x \sim_1^{+} z$ and $z \sim_2^{+} y$, so $x \simeq^{+} y$. Thus $\simeq$ is a bisimulation, and $\approx$ is transitive.

Using stratified Comprehension on $\mathcal{P}(\operatorname{Dom} R)$, we get the set A of $\approx$-equivalence classes. For convenience, write $\hat{x}$ for the equivalence class of x.

Define a stratified relation Q on A by $aQb \Leftrightarrow_{df} (\exists x \in a)(\exists y \in b)xRy$.

The set $\{x \in \operatorname{Dom} R : \hat{x} \in \hat{r}_Q\}$ contains r and is closed under R^{-1}, so it is the whole of $r_R = \operatorname{Dom} R$ and thus $\hat{r}_Q = A$.

Let $\sim_Q$ be a bisimulation on Q. Define a stratified relation on R by $x \sim_R y \Leftrightarrow_{df} \hat{x} \sim_Q \hat{y}$.

Suppose $\hat{x} \sim_Q \hat{y}$, then

$$(\forall aQ\hat{x})(\exists bQ\hat{y})(a \sim_Q b) \wedge (\forall bQ\hat{y})(\exists aQ\hat{x})(a \sim_Q b)$$

Since each $a \in \operatorname{Dom} Q$ is $\hat{z}$ for some $z \in \operatorname{Dom} R$

$$(\forall zRx)(\exists tRy)(z \sim_R t) \wedge (\forall tRy)(\exists zRx)(z \sim_R t)$$

This shows $\sim_R$ is a bisimulation on R, so $x \approx y$ and $\hat{x} = \hat{y}$. Thus any bisimulation on Q is the identity, i.e. $[Q, \hat{r}]$ is an RPG.

Therefore there is a bijection $\theta : B \leftrightarrow A$ where B is a set of singletons.

Define a relation S on $\bigcup B$ by $xSy \Leftrightarrow_{df} \theta\{x\}\ Q\ \theta\{y\}$ and let $\{s\} = \theta^{-1}\hat{r}$, then clearly $[S, s]$ is also an RPG.

Define the map $\pi : \operatorname{Dom} R \to \operatorname{Dom} S$ by stratified comprehension as

$$\{\langle x, y\rangle : (\exists a \in A)(x \in a \wedge \theta(a) = \{y\})\}$$

Clearly π is surjective and for all $z, t \in \operatorname{Dom} S$

$$zSt \Leftrightarrow (\exists x, y \in \operatorname{Dom} R)(z = \pi(x) \wedge t = \pi(y) \wedge xRy)$$

Now let $p \in \operatorname{Dom} R$ such that $[R_r, p]$ is an RPG.

The restriction of $\approx$ to p_R is a bisimulation on R_p, so it is the identity. But $\pi(x) = \pi(y) \Leftrightarrow x \approx y$, so π is injective on p_R and thus an isomorphism between R_p and $S_{\pi(p)}$. $\square$

Thus all the axioms of ZF hold in $\mathcal{R}_M$, except Infinity and Foundation.

Lemma 39. *Infinity holds in $\mathcal{R}_M$.*

Proof. As before, let R be a well-order in M with no maximal element. Let $s \notin \operatorname{Dom} R$ and define $S := R \cup (\operatorname{Dom} R \times \{s\})$. Then $[S, s]$ is an BFEXT. But all BFEXTs are RPGs as we will show in Lemma 42, so $[S, s] \in \mathcal{R}_M$. Now $[S, s]$ is well-ordered by $\lhd$ with no maximal element, so Infinity holds. $\square$

The next result allows us to prove the anti-foundation axiom for this model.

Lemma 40. *Let $\equiv$ be the relation defined in Section 2.1 to interpret the symbol $\in$ in the language of set theory (see page 26).*

If $[A, a]$ is an APG, then there exists a function $\phi : \operatorname{Dom} A \to \mathcal{R}_M$ such that

$$(\forall x \in \operatorname{Dom} A)(\forall y \in \mathcal{R}_M)(y \lhd \phi(x) \Leftrightarrow (\exists z A x) y \equiv \phi(z))$$

Furthermore, if ψ is another function with the property above, then $\phi(x) \equiv \psi(x)$ for all $x \in \operatorname{Dom} A$.

Proof. By the Axiom of Quotient, there is a quotient map π from $[A, a]$ to some RPG $[Q, q]$.

Then there is a bijection $\theta : \operatorname{Dom} Q \leftrightarrow \iota^5 “G$ for some $G \in M$. Define

$$S := \{\langle x, y\rangle : \langle \theta^{-1}\iota^5 x, \theta^{-1}\iota^5 y\rangle \in Q\}$$

By stratified Comprehension, let

$$\phi(x) := [S_y, y] \Leftrightarrow \theta\pi(x) = \iota^5 y$$

Let $x \in \operatorname{Dom} A$ and $\phi(x) = [S_y, y]$, then for any $z \in \mathcal{R}_M$

$$\begin{aligned} z \lhd [S_y, y] &\Leftrightarrow (\exists p S y) z \equiv [S_p, p] \\ &\Leftrightarrow (\exists r A x) z \equiv \phi(r) \end{aligned}$$

Now let ψ satisfy the same condition, and $[T,t] := \psi(a)$.

The set $\{x \in \operatorname{Dom} A : (\exists y \in \operatorname{Dom} T)\psi(x) \equiv [T_y, y]\}$ contains a and is closed under A^{-1} by the hypothesis, so it is $\operatorname{Dom} A$.

Define a relation $\sim$ on A by $x \sim y \Leftrightarrow_{df} \psi(x) \equiv \psi(y)$, then $\sim$ is a bisimulation on A by the hypothesis. Hence $\{\langle \phi(x), \phi(y)\rangle : x \sim y\}$ is a bisimulation on the rigid relation S, so $\psi(x) \equiv \psi(y) \Leftrightarrow \phi(x) = \phi(y)$.

Since T is also rigid, there is a bijection $\tau : \operatorname{Dom} S \leftrightarrow \operatorname{Dom} T$ given by

$$\tau(x) := y \Leftrightarrow (\exists z \in \operatorname{Dom} A)(\phi(z) = [S_x, x] \wedge \psi(z) \equiv [T_y, y])$$

But for $x, y \in \operatorname{Dom} S$

$$\begin{aligned} xSy &\Leftrightarrow (\exists zAt)(\phi(z) = [S_x, x] \wedge \phi(t) = [S_y, y]) \\ &\Leftrightarrow (\exists zAt)([T_\tau(x), \tau(x)] \equiv \psi(z) \lhd \psi(t) \equiv [T_\tau(y), \tau(y)]) \\ &\Leftrightarrow \tau(x) T \tau(y) \end{aligned}$$

Therefore τ is an isomorphism from S to T, and thus $\phi(x) \equiv \psi(x)$ for all $x \in \operatorname{Dom} A$. $\square$

Lemma 41. *AFA holds in $\mathcal{R}_M$.*

Proof. Suppose $\mathcal{R}_M$ believes $[A,a]$ is an APG with graph relation $[G,g]$ on domain $[D,d]$ and distinguished point $[D_p, p]$ where pDd. Define a graph on $\{[D_x, x] : xDd\}$ as follows

$$H := \{\langle [D_x, x], [D_y, y]\rangle : \mathcal{R}_M \models \langle [D_x, x], [D_y, y]\rangle \in [G, g]\}$$

Then $[H, [D_p, p]]$ is an APG since any set containing $[D_p, p]$ and closed under H^{-1} contains the whole of $\{[D_x, x] : xDd\}$.

Thus by the previous Lemma there is a map $\phi : \operatorname{Dom} H \to \mathcal{R}_M$, unique up to equivalence under $\equiv$, such that

$$(\forall x \in \operatorname{Dom} H)(\forall y \in \mathcal{R}_M)(y \lhd \phi(x) \Leftrightarrow (\exists z H x) y \equiv \phi(z))$$

By the Supertransitivity Lemma, there is an RPG $[S, s]$ such that for any RPG $[T, t]$

$$[T,t] \lhd [S,s] \Leftrightarrow [T,t] \in \{[R,r] : (\exists x, y \in \mathcal{R}_M)(\langle x, y\rangle \in \phi \wedge \mathcal{R}_M \models [T,t] = \langle x, y\rangle)$$

Then $\mathcal{R}_M$ believes $[S,s]$ is the function that sends $x \in \operatorname{Dom} H$ to $\phi(x)$, which is a decoration of $[A,a]$.

Conversely, if $\mathcal{R}_M$ thinks $[Q,q]$ is a decoration of $[A,a]$, then consider

$$\psi := \Big\{\langle x, y\rangle : x \in \operatorname{Dom} H \wedge y \in \{[S_x, x] : x \in \operatorname{Dom} S\} \wedge \mathcal{R}_M \models \langle x, y\rangle \in [S,s]\Big\}$$

Then ψ satisfies the same property as ϕ, so by uniqueness $\psi(x) \equiv \phi(x)$ for all $x \in \operatorname{Dom} H$. Therefore $[Q, q] \equiv [S, s]$, i.e. the decoration is unique. $\square$

This completes the proof of:

Theorem 3. *If M is a model of strZF$_{RPG}$, then the class of RPGs in M is a model of ZFA.*

2.4 Relation between the $\mathcal{R}_M$ and $\mathcal{B}_M$

Lemma 42. *BFEXTs are exactly the same as well-founded RPGs.*

Proof. Let $[R, r]$ be a BFEXT and $\sim$ be a bisimulation on R.

Let x be R-minimal in the set $\{x \in \operatorname{Dom} R : (\exists y \neq x) x \sim y\}$ and y be R-minimal corresponding to x. Then

$$(\forall z R x)(\forall t R y)(z \sim t \Rightarrow z = t)$$

But $\sim$ is a bisimulation so $\{z : zRx\} = \{t : tRy\}$, thus $x = y$ by Extensionality. Contradiction, hence $[R, r]$ is an RPG.

Conversely, let $[R, r]$ be an RPG where R is well-founded.

Let $x, y \in \operatorname{Dom} R$ be such that $\{z : zRx\} = \{z : zRy\}$, then the map $x \mapsto y$, $z \mapsto z$ for $z \neq x, y$ is a bisimulation on R. Hence $x = y$, so $[R, r]$ is a BFEXT. $\square$

Thus if M is a model of strZF$_{RPG}$, then we have both a model of ZFA in $\mathcal{R}_M$ and a model of ZF in $\mathcal{B}_M$.

Lemma 43. *Let $[R, r]$ be an RPG. Then $\mathcal{R}_M$ thinks $[R, r]$ is well-founded if and only if R is well-founded.*

Proof. Suppose $\mathcal{R}_M$ thinks $[R, r]$ is well-founded. Let $A \subset \operatorname{Dom} R$, then by the Supertransitivity Lemma we have an RPG $[S, s]$ such that

$$(\forall [P, p] \in \mathcal{R}_M)([P, p] \lhd [S, s] \Leftrightarrow (\exists x \in A)[P, p] \equiv [R_x, x])$$

Then $\mathcal{R}_M \models [S, s] \subset [R, r]$, so let $x \in A$ be such that $\mathcal{R}_M$ thinks $[R_x, x]$ is the minimal member of $[S, s]$. But then x is R-minimal in A, otherwise there would be $y \in A$ such that $[R_y, y] \lhd [R_x, x]$.

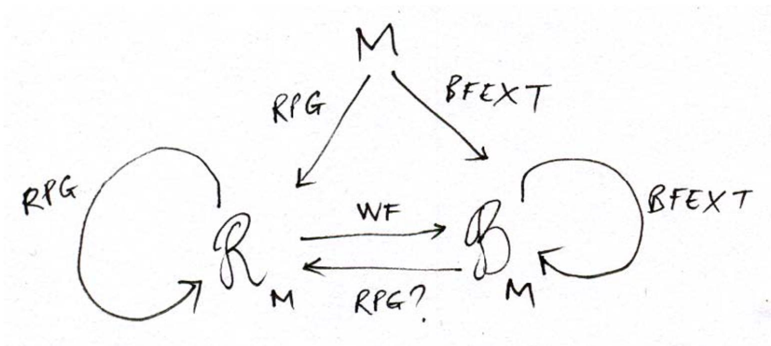
Conversely, let R be well-founded. Let $\mathcal{R}_M$ think $[S, s] \subset [R, r]$, and $A := \{xRr : [R_x, x] \lhd [S, s]\}$. Let x be R-minimal in A, then $[R_x, x]$ is $\lhd$-minimal in $[S, s]$. $\square$

Remark 27. *Thus we easily see that $\mathcal{B}_M$ is precisely the well-founded part of $\mathcal{R}_M$.*

The exact isomorphism cannot be seen from within M since the equivalence classes of graphs are too large; but as in the case of V and $\mathcal{B}_{HSS}$ with Choice, first-order elementary equivalence can still be proven (cf. Remark 20).

Remark 28. *Since $\mathcal{B}_M$ is a model of ZF, if we apply the BFEXT construction again, we will just get the same model - since every BFEXT in it can be Mostowski collapsed down and every transitive membership graph is automatically an BFEXT. Similarly, for $\mathcal{R}_M$ the anti-foundation axiom means that applying the RPG construction again gives us the same model back.*

We know that taking the internal well-founded part of $\mathcal{R}_M$ gives $\mathcal{B}_M$ - so what does the RPG construcion on $\mathcal{B}_M$ look like? If $M = HSS$ and we have the conditions that lead to $\mathcal{B}_M \cong V$, then the RPG construction on $\mathcal{B}_M$ is the same as that on V, and isomorphic to $\mathcal{R}_M$. If Choice does not hold, when do we get $\mathcal{R}_M$?



2.5 Application to hereditarily symmetric sets

We show that HSS is a model of strZF$_{RPG}$ if $G \cap HSS \in \mathcal{F}$, hence the class of RPGs in HSS models ZFA and the class of BFEXTs in HSS models ZF. Since $G \cap HSS \in \mathcal{F}$ implies the strong closure condition, HSS is already a model of strZF. Thus it only remains to show Weak IO for RPGs.

Lemma 44. *If $[R, r]$ is an RPG in HSS where $G \cap HSS \in \mathcal{F}$, then* $\operatorname{Dom} R$ *is uniformly symmetric.*

Proof. Let $R \in HSS$ and $G_n(R) \in \mathcal{F}$. Let $H := G_n(R) \cap HSS$, then $H \in \mathcal{F}$.

If $\sigma \in H$ and $i \geq n$, then $j^i\sigma$ is a permutation on $\operatorname{Dom} R$ and its graph on $\operatorname{Dom} R$ is in HSS since $\sigma \in HSS$.

If $\langle x, y\rangle \in R$, then $\langle j^i\sigma(x), j^i\sigma(y)\rangle = j^{i+2}\sigma\langle x, y\rangle \in R$. Hence $j^i\sigma$ is an automorphism of R in HSS. By the Axiom of Stability for RPGs (which can be proved in strZF$^-$) $j^i\sigma$ is the identity on $\mathrm{Dom}\, R$.

Then

$$(\forall x \in \mathrm{Dom}\, R)(\forall \sigma \in H)(\forall i \geq n+3) j^i\sigma(x) = x$$

so $\mathrm{Dom}\, R$ is uniformly symmetric. □

Remark 29. *An easy adaptation of the proof above shows the same result for HWS.*

Corollary 8. *Weak IO for RPGs holds in HSS.*

Proof. Every uniformly symmetric set is strongly Cantorian. □

Even if $G \cap HSS \notin \mathcal{F}$, we can still prove Weak IO for BFEXT in HSS.

Lemma 45. *If $[R, r]$ is a BFEXT in HSS, then* $\mathrm{Dom}\, R$ *is uniformly symmetric.*

Proof. Let $G_n(R) \in \mathcal{F}$. As before if $\sigma \in G_n(R)$ and $i \geq n$, then $j^i\sigma$ is an automorphism of R in V even though the graph of $j^i\sigma$ on $\mathrm{Dom}\, R$ might not be in HSS. However we already know that well-founded extensional relations in the sense HSS are well-founded extensional in V, so $j^i\sigma$ still has to be the identity on $\mathrm{Dom}\, R$. □

Thus if the strong closure condition holds but $G \cap HSS \notin \mathcal{F}$, HSS may not be a model of strZF$_{RPG}$ but it is still a model of strZF$_{BFEXT}$. Therefore even though the RPG model of ZFA may not exist, we can still build the BFEXT model of ZF from HSS.

2.5.1 Without the closure condition

What if the closure condition does not hold? Then HSS might not even satisfy Power Set, let alone strZF$_{BFEXT}$. However we can still show that the BFEXTs in HSS form a model of Zermelo set theory minus Infinity by exploiting specific properties of HSS.

Let $\mathcal{B}$ be the class of BFEXTs in V and $\mathcal{B}_{HSS}$ be the class of BFEXTs in HSS. Our previous results shows that $\mathcal{B}_{HSS} = \mathcal{B} \cap HSS$, and we will make full use of the fact that $\mathcal{B}$ is a model of ZF (under the intended interpretation of the language of set theory).

Moreover if $[R, r], [S, s] \in \mathcal{B}_{HSS}$, then $\mathrm{Dom}\, R$ and $\mathrm{Dom}\, S$ are uniformly symmetric so any function between them is uniformly symmetric as well. In particular if $[R, r] \cong [S, s]$ in V, then the isomorphism is also in HSS. This also shows $[R, r] \lhd [S, s]$ in $\mathcal{B}$ if and only if $[R, r] \lhd [S, s]$ in $\mathcal{B}_{HSS}$, so from now on we no longer need to specify context for $\cong$ or $\lhd$.

Remark 30. *Let $[R, r] \in \mathcal{B}_{HSS}$, then every $\mathcal{B}$-member of $[R, r]$ is isomorphic to a $\mathcal{B}_{HSS}$-member of $[R, r]$. To see this, note that if $[P, p] \lhd [R, r]$ in $\mathcal{B}$, then $[P, p] \cong [R_x, x]$ for some xRr.*

Lemma 46. *Extensionality holds in $\mathcal{B}_{HSS}$.*

Proof. Suppose $[R, r], [S, s] \in \mathcal{B}_{HSS}$ have the same $\mathcal{B}_{HSS}$-members. By Remark 30 $[R, r]$ and $[S, s]$ must have the same $\mathcal{B}$-members, so they are isomorphic by Extensionality in $\mathcal{B}$. □

Remark 31. *Let $[R, r]$ be an EPG made of BFEXTs in $\mathcal{B}$ and $[Q, q] \in \mathcal{B}$ the quotient of $[R, r]$ as given by the Axiom of Quotient. If $R \in HSS$ and is uniformly symmetric, an inspection of the proof of the Axiom of Quotient for BFEXTs shows that $[Q, q] \in HSS$.*

Definition 21. *Say $A \subset \mathcal{B}_{HSS}$ is* supersymmetric *if there exists $n \in \omega$ and $H \in \mathcal{F}$ such that*

$$(\forall [R, r] \in A)(\forall x \in \operatorname{Dom} R) H \subset G_n(x)$$

Remark 32. *Let $A \subset \mathcal{B}_{HSS}$ be supersymmetric. By the Supertransitivity Lemma for $\mathcal{B}$ there exists $[Q, q] \in \mathcal{B}$ such that*

$$(\forall [P, p] \in \mathcal{B})([P, p] \lhd [Q, q] \Leftrightarrow [P, p] \in A$$

Furthermore in the proof of the Supertransitivity Lemma (Lemma 25) if the new vertex $b \in HSS$, one can easily check that by supersymmetry of A the graph $B \in HSS$ and is uniformly symmetric. Then $Q \in HSS$ by Remark 31 so $[Q, q] \in \mathcal{B}_{HSS}$.

Lemma 47. *Comprehension holds in $\mathcal{B}_{HSS}$.*

Proof. Let $[R, r] \in \mathcal{B}_{HSS}$ and $\phi(x)$ be a formula with parameters in $\mathcal{B}_{HSS}$. The set $\{[R_x, x] : xRr \wedge \mathcal{B}_{HSS} \models \phi([R_x, x])\}$ is supersymmetric, so by Remark 32 there exists $[Q, q] \in \mathcal{B}_{HSS}$ such that

$$(\forall [P, p] \in \mathcal{B})\Big([P, p] \lhd [Q, q] \leftrightarrow \mathcal{B}_{HSS} \models \phi([P, p])\Big)$$

By Remark 30 it is clear that $\mathcal{B}_{HSS} \models [Q, q] = \{x \in [R, r] : \mathcal{B}_{HSS} \models \phi(x)\}$. □

Lemma 48. *Pair Set holds in $\mathcal{B}_{HSS}$.*

Proof. If $[R, r], [S, s] \in \mathcal{B}_{HSS}$, it is straightforward to verify that the set $\{[R, r], [S, s]\}$ is supersymmetric. Then by Remark 32 we have the result. □

Lemma 49. *Sumset holds in $\mathcal{B}_{HSS}$.*

Proof. If $[R, r] \in \mathcal{B}_{HSS}$, the set $\{[R_x, x] : (\exists y)(xRy \wedge yRr)\}$ is supersymmetric so the result holds by Remark 32. $\square$

Lemma 50. *Power Set holds in $\mathcal{B}_{HSS}$.*

Proof. If $[R, r] \in \mathcal{B}_{HSS}$, let $\mathcal{B} \models [Q, q] = \mathcal{P}[R, r]$. Inspecting the earlier proof of Power Set for our syntactic models shows that by choosing $c \in HSS$ we can arrange that the EPG $[C, c] \in HSS$ and is uniformly symmetric. Hence $[Q, q] \in HSS$ by Remark 31 and it is clear that $\mathcal{B}_{HSS} \models [Q, q] = \mathcal{P}[R, r]$. $\square$

Remark 33. *The truth of Infinity in $\mathcal{B}_{HSS}$ depends on Infinity in HSS: given any well-order of limit type in HSS we can easily build one in $\mathcal{B}_{HSS}$ and vice versa.*

On the other hand the proof of Replacement in $\mathcal{B}_{HSS}$ runs into a different difficulty. If $[R, r] \in \mathcal{B}_{HSS}$ and $\mathcal{B}_{HSS}$ believes $\phi(x, y)$ defines a function on $[R, r]$, we would need a supersymmetric set A such that

$$(\forall x R r)(\exists [Q, q] \in A)\mathcal{B}_{HSS} \models \phi([R_x, x], [Q, q])$$

But while the set $\{[R_x, x] : xRr\}$ is easily supersymmetric, it is not clear that we could prove the same for A since the function described by ϕ could change the degree of symmetry of each $[R_x, x]$ differently.

Thus the question remains of whether there is a pair $(G, \mathcal{F})$ such that the closure condition fails for HSS and either Infinity or Replacement fails for $\mathcal{B}_{HSS}$.

3 Multisets

It has been observed by Thomas Forster that in working with stratified set theories, one frequently has difficulties with making copies of sets. If we want the copies to be indexed by some set I, the traditional approach would be to make ordered pairs with the index component being members of I. Then by reading off this component one would have a bijection between the set of copies and the index set I. However it is clear that this bijection may well be unstratified unless a type-level definition of ordered pairs is used. In the previous chapter we met some instances of this: for example in the proof of the Supertransitivity Lemma (Lemma 25) merely making disjoint copies of a given set of APGs raises their type by 5. In that particular case we were saved by Weakened IO, which let us bring the type back down.

Things get even more complicated when the index set is not disjoint from the sets being copied, in which case even type-level ordered pairs might not help. Consider the following example: Suppose we tried to interpret NF syntactically from the APGs in an NFU model in the same manner as in the previous chapter. As before the definition of APGs is stratified and the definition of a bisimulation can be carried over to enforce extensionality between equivalence classes of APGs. It is also relatively straightforward to verify Comprehension for this interpretation, and we are left with constructing a universal APG. Conveniently in the original NFU model there exists a set of all APGs so it seems we need only pick a new vertex and connect it to the top of all APGs to get the universal APG. The only thing left is to arrange for all the children APGs to have disjoint domains, so that when we join them in the universal APG they do not receive extra edges. However it is impossible to make disjoint copies of all APGs: if $\hat{A}$ is the new copy of A, we would have to index each vertex of $\hat{A}$ with A itself to be sure that $\operatorname{Dom}\hat{A}$ is disjoint from other $\operatorname{Dom}\hat{B}$. But then $\hat{A}$ will be of higher type than A even if our ordered pairs are type-level, since its vertices are already of the same type as A; so we will not have a bijection from the set of copies to the original set of APGs!

Thus in stratified theories, even something as simple as copying sets may prove troublesome. One way to try to get around this difficulty is to introduce an extra degree of freedom into our objects, by considering multisets in place of sets. Since each member of a multiset comes with a multiplicity, we can encode more information in a multiset than in a pure set, which might give a means to index things. Another reason to introduce this extra freedom from the NF perspective is that, according to a view attributed to Randall Holmes, the increased "slop" may potentially make it easier to prove the consistency of NF-like systems. To that end, we will investigate a new axiomatisation of multisets customised to be fitted with a stratification system.

Multisets are sets with possibly repeated elements, somewhat natural objects that arise in various situations in both mathematics and computing. However despite numerous accounts of multisets, some by quite well-known mathematicians, there has been no consensus on how to axiomatise them. The best survey of these accounts is Wayne Blizard's *The development of multiset theory* [Blizard 1] and the two most comprehensive proposals seem to be Blizard's own in [Blizard 2] and [Blizard 3], the latter even allowing for infinite multiplicities. However like other multiset theories, they are both two-sorted theories where the multiplicities are a different type of objects from the multisets they support. This would require separate axioms for multiplicity arithmetic, and in the infinite case it involves piggybacking on a predefined model of cardinal arithmetic (for example [Blizard 3] uses cardinals in a model of ZF set theory).

The above would not serve our purpose, since beyond the cumbersome axiomatisation the space of cardinal multiplicities may not be rich enough to encode information about the multisets themselves. Therefore we will now propose a one-sorted account of multisets, where multiplicities and sets come from the same universe and follow the same axioms. As a result multiplicities are no longer cardinal numbers but sets themselves, with their own internal structures. The natural ordering of multiplicities will be identified with the subset relation, i.e. intuitively we consider x to be less than y as multiplicities if x is a proper subset of y. The axioms we propose will mirror Zermelo-Fraenkel set theory, *mutatis mutandis*, the only real complication coming from the subset relation for multisets, which becomes naturally recursive upon being identified with the ordering on multiplicities. As an intended tool to help prove the consistency of fragments of NF, we will keep our multiset theory relatively consistent with ZF. We will also define a stratification system that gives rise to a multiset analogue of Coret's lemma and a parallel structure to the hereditarily symmetric sets.

3.1 The theory

This is a one-sorted theory where the same variables will be used for multisets and multiplicities; we shall often call them sets for brevity. The membership predicate is a ternary predicate: $x \in_a y$ means x belongs to y with multiplicity a. Note that a is of the same type of object as x and y, and in turn may have members of its own.

Definition 22. *The language of multisets $\mathcal{L}_{\mathcal{H}}$ has one sort of variable and two predicate symbols: the identity $=$ and the ternary symbol $\in$ (in practice we write $x \in_a y$).*

Remark 34. *The membership and subset relations we will define on multisets will be* *denoted by boldface symbols to differentiate them from their set-theoretic counterparts. This distinction is especially important when we build a model for our multiset theory from a model of set theory. Unfortunately the bold symbols look almost the same as the regular symbols, so the reader should keep in mind the context where they appear.*

We write $x \mathrel{\overline{\in}} y$ for $(\exists a)x \in_a y$, i.e. x belongs to y with some multiplicity.

To begin, we assert that multiplicities are unique.

Axiom 7.

$$(\forall x, y, a, b)(x \in_a y \wedge x \in_b y \Rightarrow a = b)$$

Thus we can write informally $\frac{y}{x}$ for the unique multiplicity of x in y.

Axiom 8. *Axiom of Extensionality*

$$(\forall x, y)\Big(x = y \Leftrightarrow (\forall a, b)(a \in_b x \Leftrightarrow a \in_b y)\Big)$$

Definition 23. *Let $\phi(x, y)$ be a formula with two free variables and possibly parameters such that $(\forall x)(\exists! y)\phi(x, y)$. We say ϕ defines a function-class on multisets.*

Definition 24. *Let ϕ define a function-class on multisets. We write $\{x \otimes y : \phi(x, y)\}$ for the multiset a satisfying*

$$(\forall x, y)(x \in_y a \Leftrightarrow \phi(x, y))$$

i.e. a contains x with multiplicity y if and only if $\phi(x, y)$ holds.

As a special case, for any concrete natural number $\mathbf{n}$ we write $\{x_1 \otimes y_1, \ldots x_{\mathbf{n}} \otimes y_{\mathbf{n}}\}$ for the multiset a satisfying

$$(\forall x, y)(x \in_y a \Leftrightarrow \langle x, y\rangle = \langle x_1, y_1\rangle \vee \ldots \vee \langle x, y\rangle = \langle x_{\mathbf{n}}, y_{\mathbf{n}}\rangle)$$

The multisets specified in the definition above are unique by Extensionality.

Axiom 9. *Axiom of Empty Set*

$$(\exists x)(\forall y) y \mathrel{\overline{\notin}} x$$

As usual, extensionality ensures that the empty multiset, which we denote by $\boldsymbol{\emptyset}$, is unique.

Axiom Schema 10. *Axiom schema of Comprehension*

$$(\forall x)(\exists y) y = \{a \otimes b : a \in_b x \wedge \phi(a, b)\}$$

for all formula ϕ with two free variables and possibly parameters.

Note that the set given by Comprehension inherits the multiplicities from the original set. At this stage we will avoid changing multiplicities in our axioms as much as possible. By introducing axioms that deal with multiplicities separately, we can extend the basic theory to incorporate different systems of multiset, for example those with only finite multiplicities or cardinal multiplicities.

Axiom 11. *Axiom of Pairing*

$$(\forall x, y)(\exists a)a = \{x \otimes \emptyset, y \otimes \emptyset\}$$

(see the special case of Definition 24)

We define ordered pairs in the usual manner and the Axiom of Pairing ensures their existence.

Definition 25. $\langle x, y\rangle := \Big\{\{x \otimes \emptyset\} \otimes \emptyset, \{x \otimes \emptyset, y \otimes \emptyset\} \otimes \emptyset\Big\}$

3.1.1 The subset relation

Intuitively in our theory there should be an ordering on multiplicities (namely "there are more copies of something than of another"). With this in mind, we regard a multiset x as a subset of y if and only if every member of x appears in y with greater or equal multiplicity. As mentioned before, identifying the subset relation with the ordering on multiplicities gives us a natural recursive definition of subset.

Axiom Schema 12. $\overline{\subset}$ *is a partial order and the largest class relation such that*

$$(\forall x, y)\left(x \mathrel{\overline{\subset}} y \Leftrightarrow (\forall a \mathrel{\overline{\in}} x)\left(a \mathrel{\overline{\in}} y \wedge \tfrac{x}{a} \mathrel{\overline{\subset}} \tfrac{y}{a}\right)\right)$$

In other words if $\phi(x, y)$ *is a formula (possibly with parameters) such that*

$$(\forall x, y)\Big(\phi(x, y) \Leftrightarrow (\forall a \mathrel{\overline{\in}} x)\left(a \mathrel{\overline{\in}} y \wedge \phi(\tfrac{x}{a}, \tfrac{y}{a})\right)\Big)$$

then $(\forall x, y)(\phi(x, y) \Rightarrow x \mathrel{\overline{\subset}} y)$.

Instead of adopting this axiom schema right away, we will prove it from the rest of the axioms by defining the subset relation in terms of the multi-membership relation. However at this stage we can check that these properties makes the empty multiset $\emptyset$ a subset of everything, and that $\{x \otimes \emptyset\}$ behaves like the traditional singleton in set theory.

Formally we write $x \mathrel{\overline{\subset}} y$ as shorthand for the following formula:

Definition 26.

$$x \mathrel{\overline{\subset}} y \Leftrightarrow_{df} x = y \vee$$
$$(\exists R)\Big(\langle x, y\rangle \mathrel{\overline{\in}} R \wedge (\forall v, w)\Big(\langle v, w\rangle \mathrel{\overline{\in}} R \Rightarrow (\forall a \mathrel{\overline{\in}} v)(a \mathrel{\overline{\in}} w \wedge \langle \tfrac{v}{a}, \tfrac{w}{a}\rangle \mathrel{\overline{\in}} R)\Big)\Big)$$

Lemma 51. $(\forall x)(\forall y)(x \mathrel{\overline{\in}} y \Leftrightarrow \{x \otimes \emptyset\} \mathrel{\overline{\subset}} y)$

Proof. In one direction $\{x \otimes \emptyset\} \mathrel{\overline{\subset}} y$ since $\emptyset$ is a subset of everything, while the other direction is trivial. $\square$

As part of our axioms we stipulate that $\overline{\subset}$ is antisymmetric. This is the only part of the Axiom of Subset that does not follow from the other axioms with the definition above, and we will prove this independence result in the last section of this thesis.

Axiom 13. *Axiom of Subset*

$$(\forall x, y)(x \mathrel{\overline{\subset}} y \wedge y \mathrel{\overline{\subset}} x \Rightarrow x = y)$$

It trivially follows from the definition that if $x \mathrel{\overline{\subset}} y$, then every member of x is a member of y. Furthermore $\emptyset$ is a subset of everything.

Remark 35. *At this point one may question the necessity of the axiom of Subset in our theory. After all the subset relation in set theory is trivially antisymmetric, and the recursive property of $\overline{\subset}$ makes it easy to prove antisymmetry if we have a reasonable definition of well-foundedness in the model. However in the last section we will show that the axiom of Subset is independent from the remaining axioms by means of a syntactic model.*

We move on to the definition of union. In set theory $\bigcup x$ is simply defined as the set of members of members of x since there are no worries about multiplicities, but on reflection the really useful feature of $\bigcup x$ is the fact that it is the minimal superset of all members of x. In the context of two-sorted multiset theory, this definition of union means taking the supremum of multiplicities of the same object, as opposed to what [Blizard 2] calls the *additive union* where multiplicities are added.

Axiom 14. *Axiom of Union*

$$(\forall x)(\exists b)(\forall a)(b \mathrel{\overline{\subset}} a \Leftrightarrow (\forall y \mathrel{\overline{\in}} x) y \mathrel{\overline{\subset}} a)$$

Following set theory convention, we denote the union of x by $\bigcup x$ and write $x \cup y$ for $\bigcup\{x \otimes \emptyset, y \otimes \emptyset\}$. Since $\overline{\subset}$ is antisymmetric, $\bigcup x$ is unique for every x.

We follow the same approach in defining Replacement. Let the formula ϕ define a function on x and a be in the image of x. If a is the image of more than one $y \in x$, the multiplicity of a in the new multiset (given by Replacement) should be at least as large as the multiplicity of any preimage of a in x. Thus we simply let the multiplicity of a be the $\overline{\subset}$-least upper bound (i.e. the union) of $\frac{x}{y}$ for all preimages y of a. If y is unique, it easily follows that the multiplicity of a is $\frac{x}{y}$ since $\overline{\subset}$ is both reflexive and antisymmetric.

Axiom Schema 15. *Axiom schema of Replacement*

$$(\forall x)\Bigg((\forall a \mathrel{\overline{\in}} x)(\exists! b)\phi(a,b) \Rightarrow$$
$$(\exists y)(\forall b,d)\Big(b \in_d y \Leftrightarrow (\exists a \mathrel{\overline{\in}} x)\phi(a,b) \wedge (\forall e)\Big(d \mathrel{\overline{\subset}} e \Leftrightarrow (\forall a \mathrel{\overline{\in}} x)(\phi(a,b) \Rightarrow \tfrac{x}{a} \mathrel{\overline{\subset}} e)\Big)\Big)\Bigg)$$

for all formulae ϕ with two free variables and possibly with parameters.

It is clear that the set given by Replacement is unique for each x and each function-class ϕ, and we will denote it by $Rep_\phi x$.

Lemma 52. $(\forall x, b)\Big(b = \bigcup x \Rightarrow (\forall a)(a \mathrel{\overline{\in}} b \Leftrightarrow (\exists y \mathrel{\overline{\in}} x) a \mathrel{\overline{\in}} y)\Big)$

Proof. If $(\forall y \mathrel{\overline{\in}} x) a \mathrel{\overline{\notin}} y$, by Comprehension let

$$z := \{v \otimes w : v \in_w b \wedge v \neq a\}$$

Let $y \mathrel{\overline{\in}} x$ and R be a witness to $y \mathrel{\overline{\subset}} b$. Define by Replacement

$$S := \{v \otimes w : (v \in_w R \wedge v \neq \langle y, b\rangle) \vee (v = \langle y, z\rangle \wedge \langle y, b\rangle \in_w R)\}$$

In other words S is obtained by replacing $\langle y, b\rangle$ in R with $\langle y, z\rangle$.

Since $a \mathrel{\overline{\notin}} y$ it is still true that

$$(\forall \langle v, w\rangle \mathrel{\overline{\in}} S)(\forall c \mathrel{\overline{\in}} v)(c \mathrel{\overline{\in}} w \wedge \langle \tfrac{v}{c}, \tfrac{w}{c}\rangle \mathrel{\overline{\in}} S)$$

Hence S is a witness to $y \mathrel{\overline{\subset}} z$ and thus $b \mathrel{\overline{\subset}} z$ by definition of union, so $a \mathrel{\overline{\notin}} b$.

Conversely let $y \mathrel{\overline{\in}} x$ and $a \mathrel{\overline{\in}} y$, then $a \mathrel{\overline{\in}} b$ since $y \mathrel{\overline{\subset}} b$ by definition. □

One may expect that the multiplicity of each $a \in \bigcup x$ is the union of multiplicities of a in all $b \in x$, but currently our axioms do not allow changing multiplicities. We will prove this result later on when the axiom schema of Multiplicity Replacement is introduced.

Given any multiset x there can be more than one multiset whose memmbers are exactly the subsets of x, so we will make a canonical choice for the power set of x by specifying all multiplicities in the power set to be $\emptyset$. Other choices are also possible, but this seems to be the simplest.

Definition 27. *The canonical power set of x is $\mathcal{P}x := \{y \otimes \emptyset : y \mathrel{\overline{\subset}} x\}$.*

Axiom 16. *Axiom of Power Set*

$$(\forall x)(\exists y) y = \mathcal{P}x$$

Lemma 53. *For any x, y there exists the product of x and y, namely*

$$x \times y := \{\langle v, w\rangle \otimes \emptyset : v \mathrel{\overline{\in}} x \wedge w \mathrel{\overline{\in}} y\}$$

Proof. By Comprehension from $\mathcal{P}^3(x \cup y)$. □

Lemma 54. *Intersection*

$$(\exists x)\phi(x) \Rightarrow (\exists b)(\forall a)\Big(a \mathrel{\overline{\subset}} b \Leftrightarrow (\forall y)(\phi(y) \Rightarrow a \mathrel{\overline{\subset}} y)\Big)$$

for any formula ϕ with one free variable and possibly parameters.

Proof. Let $\phi(x)$ hold for some x. By Comprehension from $\mathcal{P}x$ and Union we have the multiset

$$b := \bigcup\{v \otimes \emptyset : v \mathrel{\overline{\subset}} x \wedge (\forall y)(\phi(y) \Rightarrow v \mathrel{\overline{\subset}} y)\}$$

If $\phi(y)$ holds, then $v \mathrel{\overline{\subset}} y$ for any v in the union, hence $b \mathrel{\overline{\subset}} y$ by definition of union. Therefore $(\forall a \mathrel{\overline{\subset}} b) a \mathrel{\overline{\subset}} y$.

Conversely if $a \mathrel{\overline{\subset}} y$ for all y such that $\phi(y)$ holds, then $a \mathrel{\overline{\subset}} x$. Hence a is in the union, so $a \mathrel{\overline{\subset}} b$. □

For convenience we denote the intersection as defined in the lemma by $\bigcap_{\phi(x)} x$. For any given ϕ and x, the intersection is unique if it exists since $\overline{\subset}$ is antisymmetric.

Remark 36. *It is time for a short comment on $\emptyset$ as a multiplicity. Normally one would expect $x \in_{\emptyset} y$ to denote non-membership, since it fits with the intuition of x belonging to y zero times. However with our definition of the inclusion relation, equating $\emptyset$-multiplicity with non-membership would give rise to rather odd phenomena. For example there might be two non-empty multisets with exactly the same members but empty intersection:*

Suppose x and y are two non-empty disjoint multisets, which can always be arranged if the axiom of Pairing holds and the model has more than one non-empty object. Then $\{\emptyset \otimes x\}$ and $\{\emptyset \otimes y\}$ have the same member, namely $\emptyset$. However the multiplicity of $\emptyset$ in the intersection must be empty since it is a subset of both x and y, hence $\{\emptyset \otimes x\} \cap \{\emptyset \otimes y\} = \emptyset$.

Remark 37. *There is no negative membership in our theory, since $\emptyset$ is already the bottom multiplicity (recall that $\emptyset$ is a subset of everything, and we already chose to identify the ordering of multiplicities with the subset relation on multisets). For a quick overview of negative multiplicities and a theory of multisets with integer multiplicities (including the negative integers), see [Blizard 4].*

If $\phi(x)$ is the formula $x \mathrel{\overline{\in}} y$, we simply write $\bigcap y$. If $\phi(x)$ is the formula $x \mathrel{\overline{\in}} a \wedge x \mathrel{\overline{\in}} b$, we write $a \cap b$.

3.1.2 Relations and functions

Definition 28. *A multiset R is a (binary) relation if all its members are ordered pairs. Define the canonical field of R by Comprehension as*

$$\operatorname{Dom} R := \left\{ x \otimes v : (\exists y)(\langle x, y\rangle \mathrel{\overline{\in}} R \vee \langle y, x\rangle \mathrel{\overline{\in}} R) \wedge x \in_v \bigcup^2 R \right\}$$

Definition 29. *A relation f is a function if*

$$(\forall a, x, y)((\langle a, x\rangle \mathrel{\overline{\in}} f \wedge \langle a, y\rangle \mathrel{\overline{\in}} f) \Rightarrow x = y)$$

We define the canonical domain and range of f as follows

$$\operatorname{dom} f := \left\{ x \otimes v : (\exists y)\langle x, y\rangle \mathrel{\overline{\in}} f \ \wedge x \in_v \bigcup^2 R \right\}$$

$$\operatorname{ran} f := \left\{ y \otimes v : (\exists x)\langle x, y\rangle \mathrel{\overline{\in}} f \ \wedge y \in_v \bigcup^2 R \right\}$$

Note that if f is a function, then $\operatorname{Dom} f = \operatorname{dom} f \cup \operatorname{ran} f$. In general f can be regarded as a function on any multiset with the same members as $\operatorname{dom} f$.

Intuitively the definition above requires a function to send all identical copies of an object in its domain to identical images. In other words a function is just a map which sends multisets to multisets, and its domain is just a canonical object to represent the class of multisets on which the map is defined. The multiplicities in the domain and in the graph of the function itseld are thus of no importance.

We write $R \mathrel{\overline{\in}} \mathit{Relation}$ and $f \mathrel{\overline{\in}} \mathit{Function}$ as shorthand for the formulae saying R is a relation and f is a function respectively, and write $f(x) = y$ for $\langle x, y\rangle \in f$. For convenience we also extend our notation by introducing $\{f(x) \otimes g(y) : \phi(x, y)\}$ where f, g are either functions or formulae defining function-classes on multisets, with the obvious meaning.

Lemma 55. $\overline{\subset}$ *is transitive and thus a partial order.*

Proof. Suppose $a \mathrel{\overline{\subset}} b$ and $b \mathrel{\overline{\subset}} c$. If $a = b$ or $b = c$ the proof is trivial, so we only consider the case where they are distinct.

If R_1 witnesses $a \mathrel{\overline{\subset}} b$ and R_2 witnesses $b \mathrel{\overline{\subset}} c$, define by Comprehension from $\bigcup^2 R_1 \times \bigcup^2 R_2$ the following relation

$$R := \{\langle x, z\rangle \otimes v : (\exists b)(\langle x, y\rangle \mathrel{\overline{\in}} R_1 \wedge \langle y, z\rangle \mathrel{\overline{\in}} R_2) \wedge \langle x, z\rangle \in_v \operatorname{Dom} R_1 \times \operatorname{Dom} R_2\}$$

If $\langle x, z\rangle \mathrel{\overline{\in}} R$, let $\langle x, y\rangle \mathrel{\overline{\in}} R_1$ and $\langle y, z\rangle \mathrel{\overline{\in}} R_2$. Suppose $v \mathrel{\overline{\in}} x$, then $v \mathrel{\overline{\in}} y$ and so $v \mathrel{\overline{\in}} z$. Furthermore $\langle \frac{x}{v}, \frac{y}{v}\rangle \mathrel{\overline{\in}} R_1$ and $\langle \frac{y}{v}, \frac{z}{v}\rangle \mathrel{\overline{\in}} R_2$ so $\langle \frac{x}{v}, \frac{z}{v}\rangle \mathrel{\overline{\in}} R$. Thus we have

$$\langle x, y\rangle \mathrel{\overline{\in}} R \Rightarrow (\forall v \mathrel{\overline{\in}} x)(v \mathrel{\overline{\in}} y \wedge \langle \tfrac{x}{v}, \tfrac{y}{v}\rangle \mathrel{\overline{\in}} R)$$

and it is trivial to see that $\langle a, c\rangle \mathrel{\overline{\in}} R$. □

3.1.3 Well-orders and Infinity

Definition 30. *Say X is a closed multiset and write $X \mathrel{\overline{\in}} Closed$ if*

$$(\forall v \mathrel{\overline{\in}} X)(\forall a, b)(a \in_b v \Rightarrow (a \mathrel{\overline{\in}} X \wedge b \mathrel{\overline{\in}} X))$$

Well-orders are defined in exactly the same way as in set theory.

Definition 31. *A relation R is a well-order if it is a total order and any non-empty multiset $A \mathrel{\overline{\subset}} \operatorname{Dom} R$ has a R-minimal member.*

To state the Axiom of Infinity we assert the existence of an analogue of the von Neumann ordinals.

Definition 32. *Write $\alpha \mathrel{\overline{\in}} ON$ for the formula saying both of the following:*

- $(\forall x \mathrel{\overline{\in}} \alpha)\left(\frac{\alpha}{x} = \emptyset \wedge (\forall y \mathrel{\overline{\in}} x)(y \mathrel{\overline{\in}} \alpha \wedge \frac{x}{y} = \emptyset)\right)$
- *The relation $\overline{\in}$ restricted to α is a well-order.*

Write α^+ for $\alpha \cup \{\alpha \otimes \emptyset\}$, and write $\alpha < \beta$ for $\alpha \mathrel{\overline{\in}} \beta$ when α and β are both ordinals.

Lemma 56. *If $\alpha \mathrel{\overline{\in}} ON$, then α is closed. Furthermore if $\beta \mathrel{\overline{\in}} \alpha$, then $\beta \mathrel{\overline{\in}} ON$.*

Proof. By comparing the definition of ordinals with Definition 30, for α to be closed it is enough to show that $\emptyset \mathrel{\overline{\in}} a$. But if γ be the $\overline{\in}$-minimal member of α, then any member of γ would also be a member of α; so γ has to be empty.

Suppose $\beta \overline{\in} alpha$ and $x \overline{\in} \beta$, then $\frac{\beta}{x} = \mathbf{\emptyset}$ since $\alpha \overline{\in} ON$. Furthermore $x \overline{\in} \alpha$, so for any $y \overline{\in} x$ we also have $y \overline{\in} \alpha$ and $\frac{x}{y} = \mathbf{\emptyset}$. Since the relation $\overline{\in}$ is a well-order on α, either $\beta \overline{\in} y$ or $y \overline{\in} \beta$. But $\beta \overline{\in} y$ would violate well-foundedness of $\overline{\in}$ on α, so $y \overline{\in} \beta$.

Hence for β to be an ordinal it only remains to prove that the relation $\overline{\in}$ restricted to β is a well-order. But we already know that every member of β is a member of α, so the result follows trivially from the fact that $\overline{\in}$ restricted to α is a well-order. □

Remark 38. *The advantages of specifying all multiplicities to be $\mathbf{\emptyset}$ are that our later construction of a ZF model inside a model of our multiset theory is simplified, and that ordinals are automatically closed multisets; though other choices are also possible. It also follows directly from the definition of α^+ that $\alpha \in_{\mathbf{\emptyset}} \alpha^+$ (note that α cannot be a member of itself due to the well-ordering condition in the definition of ordinals).*

Just like in traditional set theory, it follows trivially that $\alpha^+ \overline{\in} ON$ whenever $\alpha \overline{\in} ON$.

Axiom 17. *Axiom of Infinity*

$$(\exists x \overline{\in} ON)\Big(\mathbf{\emptyset} \in_{\mathbf{\emptyset}} x \wedge (\forall y \overline{\in} x)\Big(y \overline{\in} ON \wedge (y = \mathbf{\emptyset} \vee (\exists z \overline{\in} x) y = z^+) \wedge y^+ \in_{\mathbf{\emptyset}} x\Big)\Big)$$

We call this multiset $\boldsymbol{\omega}$ and use the usual numerals to stand for the appropriate finite ordinals, i.e. $n+1$ stands for n^+.

Remark 39. *The usual proof of induction on the ordinals also works here: Suppose the formula $\phi(\alpha)$ (possibly with parameters) holds for an ordinal α whenever it holds for all ordinals $\beta < \alpha$. If $\phi(\gamma)$ is false for some $\gamma \in ON$, let δ be the least ordinal in γ^+ such that $\phi(\delta)$ is false (since the membership relation restricted to γ^+ is a well-order). Then ϕ holds for all ordinals smaller than δ since γ^+ is closed (i.e. any ordinal smaller than δ is also a member of γ^+), so $\phi(\delta)$ is true and we get a contradiction.*

3.1.4 The maximal property of $\overline{\subset}$

Lemma 57. *For any x, there exists a closed multiset with x as a member.*

Proof. Define a function-class

$$\varphi(v, w) \Leftrightarrow_{df} w = v \cup \{b \otimes b : (\exists a) a \in_b v\}$$

By Union and Replacement

$$(\forall v)(\exists w)\varphi(v, w)$$

Define a function-class $\phi(x, y)$ from $\boldsymbol{\omega}$ to the universe of multisets as follows

$$\begin{aligned}\phi(x, y) \Leftrightarrow_{df} (\exists f \in \mathit{Function})\Big(& f(x) = y \\ &\wedge \operatorname{dom} f \mathrel{\overline{\sqsubset}} \boldsymbol{\omega} \\ &\wedge (\forall n \in \operatorname{dom} f)(\forall m < n) m \mathrel{\overline{\in}} \operatorname{dom} f \\ &\wedge f(0) = \{x \otimes \boldsymbol{\emptyset}\} \\ &\wedge (\forall n \mathrel{\overline{\in}} \operatorname{dom} f) f(n+1) = \bigcup \mathit{Rep}_{\varphi} f(n)\Big)\end{aligned}$$

By induction on $\boldsymbol{\omega}$ (see Remark 39) it is easy to see that ϕ defines a function on all of $\boldsymbol{\omega}$, so by Replacement let $X := \mathit{Rep}_{\phi}\boldsymbol{\omega}$. Thus

$$(\forall y)(y \mathrel{\overline{\in}} X \Leftrightarrow (\exists a \mathrel{\overline{\in}} \boldsymbol{\omega})\phi(a, y))$$

For any $v \mathrel{\overline{\in}} X$, there exists $n \mathrel{\overline{\in}} \boldsymbol{\omega}$ such that $v \mathrel{\overline{\in}} f(n)$.

Let w be such that $\phi(v, w)$ holds. If $a \in_b v$, then $a, b \mathrel{\overline{\in}} w$ so $a, b \mathrel{\overline{\in}} f(n+1)$.

Thus $\bigcup X$ has the desired closure property, and $x \mathrel{\overline{\in}} \bigcup X$ since $x \mathrel{\overline{\in}} f(0)$. □

Corollary 9. *For any multiset x there is a $\overline{\sqsubset}$-minimal closed multiset with x as a member.*

Proof. Let $\phi(y)$ be the formula stating that y is closed and $x \mathrel{\overline{\in}} y$, then by the last lemma there is at least one y such that $\phi(y)$ holds. Hence we can take the intersection of all multisets satisfying ϕ. □

Lemma 58.

$$(\forall x, y)\Big(x \mathrel{\overline{\sqsubset}} y \Leftrightarrow (\forall a \mathrel{\overline{\in}} x)(a \mathrel{\overline{\in}} y \wedge \tfrac{x}{a} \mathrel{\overline{\sqsubset}} \tfrac{y}{a})\Big)$$

Moreover if $\phi(x, y)$ is a formula such that

$$(\forall x, y)\Big(\phi(x, y) \Rightarrow (\forall a \mathrel{\overline{\in}} x)(a \mathrel{\overline{\in}} y \wedge \phi(\tfrac{x}{a}, \tfrac{y}{a}))\Big)$$

then $(\forall x, y)(\phi(x, y) \Rightarrow x \mathrel{\overline{\sqsubset}} y)$.

Proof. If $x = y$ the first claim is trivial, so we assume otherwise.

Let $x \mathrel{\overline{\sqsubset}} y$, then $\langle x, y\rangle \mathrel{\overline{\in}} R$ for some relation R such that

$$(\forall v, w)\Big(\langle v, w\rangle \mathrel{\overline{\in}} R \Rightarrow (\forall a \mathrel{\overline{\in}} v)(a \mathrel{\overline{\in}} w \wedge \langle \tfrac{v}{a}, \tfrac{w}{a}\rangle \mathrel{\overline{\in}} R)\Big)$$

If $a \mathrel{\overline{\in}} x$, then $a \mathrel{\overline{\in}} y$ and $\langle \frac{x}{a}, \frac{y}{a}\rangle \mathrel{\overline{\in}} R$. Hence R itself witnesses $\frac{x}{a} \mathrel{\overline{\sqsubset}} \frac{y}{a}$.

Conversely, suppose

$$(\forall a \mathrel{\overline{\in}} x)(a \mathrel{\overline{\in}} y \wedge \tfrac{x}{a} \mathrel{\overline{\sqsubset}} \tfrac{y}{a})$$

Let X be a closed multiset containing x and Y a closed multiset containing y. Define a relation R by Comprehension from $X \times Y$ as follows

$$R := \Big\{ \langle v, w \rangle \otimes a : (\forall b \mathrel{\overline{\in}} v) \left(b \mathrel{\overline{\in}} w \wedge \tfrac{v}{b} \mathrel{\overline{\subset}} \tfrac{w}{b} \right) \wedge \langle v, w \rangle \in_a X \times Y \Big\}$$

By the proven direction of the claim

$$(\forall v \mathrel{\overline{\in}} X)(\forall w \mathrel{\overline{\in}} Y)(v \mathrel{\overline{\subset}} w \Rightarrow \langle v, w \rangle \mathrel{\overline{\in}} R)$$

From this and the definition of R we have

$$(\forall v, w)(\langle v, w \rangle \mathrel{\overline{\in}} R \Rightarrow (\forall b \mathrel{\overline{\in}} v)(b \mathrel{\overline{\in}} w \wedge \langle \tfrac{v}{b}, \tfrac{w}{b} \rangle \mathrel{\overline{\in}} R))$$

But $\langle x, y \rangle \mathrel{\overline{\in}} R$ by the hypothesis, so $x \mathrel{\overline{\subset}} y$ as witnessed by R and the converse is proved.

Now suppose $\phi(x, y)$ is a formula such that

$$(\forall x, y)\Big(\phi(x, y) \Rightarrow (\forall a \mathrel{\overline{\in}} x) \left(a \mathrel{\overline{\in}} y \wedge \phi(\tfrac{x}{a}, \tfrac{y}{a}) \right) \Big)$$

and that $\phi(x, y)$ holds for some particular pair x, y. Again, let X, Y be closed multisets containing x, y respectively and define a relation R by

$$R := \Big\{ \langle v, w \rangle \otimes a : (\forall b \mathrel{\overline{\in}} v) \left(b \mathrel{\overline{\in}} w \wedge \phi(\tfrac{v}{b}, \tfrac{w}{b}) \right) \wedge \langle v, w \rangle \in_a X \otimes Y \Big\}$$

As above we have

$$(\forall v \mathrel{\overline{\in}} X)(\forall w \mathrel{\overline{\in}} Y)(\phi(v, w) \Rightarrow \langle v, w \rangle \mathrel{\overline{\in}} R)$$

and

$$(\forall v, w)(\langle v, w \rangle \mathrel{\overline{\in}} R \Rightarrow (\forall b \mathrel{\overline{\in}} v)(b \mathrel{\overline{\in}} w \wedge \langle \tfrac{v}{b}, \tfrac{w}{b} \rangle \mathrel{\overline{\in}} R)$$

Hence $x \mathrel{\overline{\subset}} y$ as witnessed by R. $\square$

3.1.5 Transitive closures

There are two obvious candidates for the definition of transitive multisets:

$$(\forall x \mathrel{\overline{\in}} a)(\forall y \mathrel{\overline{\in}} x) y \mathrel{\overline{\in}} a$$

and

$$(\forall x \mathrel{\overline{\in}} a) x \mathrel{\overline{\subset}} a$$

The second trivially implies the first, but unlike in set theory the converse is false: let $x := \{\emptyset \otimes \{\emptyset\}\}$ and $a := \{x, \emptyset\}$, then a satisfies the first condition but not the second. Therefore we will take the stronger condition to be our definition for transitive multisets.

Definition 33. *A multiset a is transitive if $(\forall x \mathrel{\overline{\in}} a) x \mathrel{\overline{\subset}} a$.*

Lemma 59. *Let $\phi(x)$ be a formula with one free variable such that $(\exists x)\phi(x)$ and $\phi(x)$ only holds for transitive multisets. Then $\bigcap_\phi x$ is transitive.*

Proof. Let $a \mathrel{\overline{\in}} \bigcap_\phi x$ and suppose $\phi(x)$ holds, then $a \mathrel{\overline{\in}} x$ since $\bigcap_\phi x \mathrel{\overline{\subset}} x$. By transitivity $a \mathrel{\overline{\subset}} x$ for all such x, so $a \mathrel{\overline{\subset}} \bigcap_\phi x$ by definition. □

Remark 40. *According to this definition of transitivity, all ordinals are transitive. For if $\alpha \mathrel{\overline{\in}} ON$ and $x \mathrel{\overline{\in}} \alpha$, then for all $y \mathrel{\overline{\in}} x$ we also have $y \mathrel{\overline{\in}} \alpha$. But from the definition of ON we also have $\frac{x}{y} = \boldsymbol{\emptyset} \mathrel{\overline{\subset}} \frac{\alpha}{y}$, so $x \mathrel{\overline{\subset}} \alpha$.*

Lemma 60. *Transitive Closure*
For any multiset x, there exists a $\overline{\subset}$-minimal transitive multiset $TC(x)$ such that $x \mathrel{\overline{\subset}} TC(x)$.

Proof. Let x be any multiset. Define a function-class ϕ by

$$\begin{aligned}\phi(a, b) \Leftrightarrow_{df} (\exists f \mathrel{\overline{\in}} \mathit{Function})\Big(f(a) = b& \\ &\wedge \operatorname{dom} f \mathrel{\overline{\subset}} \boldsymbol{\omega} \\ &\wedge (\forall c \mathrel{\overline{\in}} \operatorname{dom} f)(\forall d < c) d \mathrel{\overline{\in}} \operatorname{dom} f \\ &\wedge f(0) = x \\ &\wedge (\forall c, d \mathrel{\overline{\in}} \operatorname{dom} f)(d = c + 1 \Rightarrow f(d) = \bigcup f(c))\Big)\end{aligned}$$

By induction we can prove that for any $a \mathrel{\overline{\in}} \boldsymbol{\omega}$ there is a unique b such that $\phi(a, b)$ holds. Hence by Replacement there exists

$$v := \{\phi(a) \otimes b : a \in_b \boldsymbol{\omega}\}$$

Let $w := \bigcup v$, then $x \mathrel{\overline{\subset}} w$ since $x \mathrel{\overline{\in}} v$.

For any $e \mathrel{\overline{\in}} w$ there exist a, b such that $\phi(a, b)$ and $e \mathrel{\overline{\in}} b$. Then $\bigcup b \mathrel{\overline{\in}} v$ by definition of ϕ, so $\bigcup b \mathrel{\overline{\subset}} w$. But $e \mathrel{\overline{\subset}} \bigcup b$, so $e \mathrel{\overline{\subset}} w$.

Thus there is a transitive w such that $x \mathrel{\overline{\subset}} w$, so let $TC(x)$ be the intersection of all such w. By the previous lemma $TC(x)$ is transitive, and by definition of intersection $x \mathrel{\overline{\subset}} TC(x)$ and $TC(x)$ is $\overline{\subset}$-minimal. □

Corollary 10.

$$(\forall x, y)(x \mathrel{\overline{\in}} TC(y) \Rightarrow TC(x) \mathrel{\overline{\subset}} TC(y)\}$$

$$(\forall x, y)\Big(x \mathrel{\overline{\in}} TC(y) \Leftrightarrow (x \mathrel{\overline{\in}} y \vee (\exists a \mathrel{\overline{\in}} y) x \mathrel{\overline{\in}} TC(a))\Big)$$

Proof. If $x \mathrel{\overline{\in}} TC(y)$, then $x \mathrel{\overline{\sqsubset}} TC(y)$ as $TC(y)$ is transitive so by definition $TC(x) \mathrel{\overline{\sqsubset}} TC(y)$. Consequently

$$(x \mathrel{\overline{\in}} y \vee (\exists a \mathrel{\overline{\in}} y) x \mathrel{\overline{\in}} TC(a)) \Rightarrow x \mathrel{\overline{\in}} TC(y)$$

Conversely suppose $x \mathrel{\overline{\notin}} y$ and $(\forall a \mathrel{\overline{\in}} y) x \mathrel{\overline{\notin}} TC(a)$. By Replacement, Comprehension and Union let

$$v := y \cup \bigcup \{TC(a) \otimes b : a \in_b y\}$$

If $w \mathrel{\overline{\in}} v$, then clearly $w \neq x$. Either $w \mathrel{\overline{\in}} y$ so $w \mathrel{\overline{\sqsubset}} TC(w) \mathrel{\overline{\sqsubset}} v$, or $w \mathrel{\overline{\in}} TC(a)$ for some $a \mathrel{\overline{\in}} y$ so $w \mathrel{\overline{\sqsubset}} TC(a) \mathrel{\overline{\sqsubset}} v$.

Hence v is transitive and $TC(y) \mathrel{\overline{\sqsubset}} v$, so $x \mathrel{\overline{\notin}} TC(y)$. □

In summary, we propose the following basic system.

Definition 34. *The theory MS consists of the following axioms:*

- *Axiom of Extensionality*

$$(\forall x, y)(x = y \Leftrightarrow (\forall a, b) a \in_b x \Leftrightarrow a \in_b y)$$

- *Axiom schema of Comprehension*

$$(\forall x)(\exists y)(\forall a, b)\Big(a \in_b y \Leftrightarrow (a \in_b x \wedge \phi(a, b))\Big)$$

  for all formula with two free variables and possibly parameters.

- *Axiom of Pairing*

$$(\forall x, y)(\exists a)\Big(x \in_{\boldsymbol{\emptyset}} a \wedge y \in_{\boldsymbol{\emptyset}} a \wedge (\forall b)(b \mathrel{\overline{\in}} a \Rightarrow (b = x \vee b = y))\Big)$$

- *Axiom of Subset*

$$(\forall x, y)(x \mathrel{\overline{\sqsubset}} y \wedge y \mathrel{\overline{\sqsubset}} x \Rightarrow x = y)$$

  where

$$x \mathrel{\overline{\sqsubset}} y \Leftrightarrow_{df} x = y \vee (\exists R)\Big(\langle x, y\rangle \mathrel{\overline{\in}} R \wedge (\forall v, w)\Big(\langle v, w\rangle \mathrel{\overline{\in}} R \Rightarrow (\forall a \mathrel{\overline{\in}} v)(a \mathrel{\overline{\in}} w \wedge \langle \tfrac{v}{a}, \tfrac{w}{a}\rangle \mathrel{\overline{\in}} R)\Big)\Big)$$

- *Axiom of Union*

$$(\forall x)(\exists b)(\forall a)(b \mathrel{\overline{\sqsubset}} a \Leftrightarrow (\forall y \mathrel{\overline{\in}} x) y \mathrel{\overline{\sqsubset}} a)$$

- *Axiom schema of Replacement*

$$(\forall x)\Big((\forall a \mathrel{\overline{\in}} x)(\exists! b)\phi(a,b) \Rightarrow$$
$$(\exists y)(\forall b,d)\Big(b \in_d y \Leftrightarrow (\exists a \mathrel{\overline{\in}} x)\phi(a,b) \wedge (\forall e)\Big(e \mathrel{\overline{\subset}} d \Leftrightarrow (\forall a \mathrel{\overline{\in}} x)(\phi(a,b) \Rightarrow \tfrac{x}{a} \mathrel{\overline{\subset}} e)\Big)\Big)\Big)$$

for all formulae ϕ with two free variables and possibly with parameters.

- *Axiom of Power Set*

$$(\forall x)(\exists y)(\forall a)\Big(a \mathrel{\overline{\in}} y \Leftrightarrow (a \mathrel{\overline{\subset}} x \wedge \tfrac{y}{a} = \emptyset)\Big)$$

- *Axiom of Infinity*

$$(\exists x \mathrel{\overline{\in}} ON)\Big(\emptyset \in_{\emptyset} x \wedge (\forall y \mathrel{\overline{\in}} x)(y \mathrel{\overline{\in}} ON \wedge y^{+} \in_{\emptyset} x \wedge (y = \emptyset \vee (\exists z \mathrel{\overline{\in}} x)y = z^{+}))\Big)$$

We leave the Axiom of Foundation for a later discussion in Section 3.1.8.

3.1.6 The collection of sets

In this section we work in a model of the theory MS.

Definition 35. *For any multiset x, let $core(x)$ be the multiset such that*

$$(\forall y)(y \mathrel{\overline{\in}} core(x) \Leftrightarrow y \mathrel{\overline{\in}} x) \wedge (\forall y \mathrel{\overline{\in}} core(x))\frac{core(x)}{y} = \emptyset$$

Let $Core$ be the class of cores, i.e. write $x \mathrel{\overline{\in}} Core$ for $(\forall y \mathrel{\overline{\in}} x)\frac{x}{y} = \emptyset$.

By Extensionality, $core(x)$ is unique if it exists. It trivially follows that the core of $core(x)$ is itself, and two multisets have the same core if and only if they have the same members. Note that if $\alpha \in ON$, then $core(\alpha) = \alpha$ by definition.

Definition 36. *We say $x \mathrel{\overline{\in}} Set$ if $x \mathrel{\overline{\in}} Core \wedge (\forall y \mathrel{\overline{\in}} TC(x))y \mathrel{\overline{\in}} Core$.*

It immediately follows from the definition that

$$(\forall x)(x \mathrel{\overline{\in}} Set \Leftrightarrow x \mathrel{\overline{\in}} Core \wedge (\forall y \mathrel{\overline{\in}} x)y \mathrel{\overline{\in}} Set)$$

Consider the following interpretation of the language of set theory:

- $=$ is the identity relation.
- $\in$ is interpreted as $\overline{\in}$.

- $(\forall x)\phi(x)$ is replaced with $(\forall x)(x \mathrel{\overline{\in}} Set \Rightarrow \phi(x))$.
- $(\exists x)\phi(x)$ is replaced with $(\exists x)(x \mathrel{\overline{\in}} Set \wedge \phi(x))$.

We first note that under this interpretation the induced subset relation coincides with $\overline{\subset}$, i.e.

$$(\forall x, y \mathrel{\overline{\in}} Set)(x \mathrel{\overline{\subset}} y \Leftrightarrow (\forall a \mathrel{\overline{\in}} x) a \mathrel{\overline{\in}} y)$$

Theorem 4. *If MS is consistent, so is ZF.*

Proof. We check that the given interpretation of the language of set theory turns Set into a model of all ZF axioms minus Foundation.

- Extensionality

  $$(\forall x, y \mathrel{\overline{\in}} Set)((\forall a \mathrel{\overline{\in}} Set)(a \mathrel{\overline{\in}} x \Leftrightarrow a \mathrel{\overline{\in}} y) \Rightarrow x = y)$$

  Suppose $a \in_b x$, then $a \mathrel{\overline{\in}} x$ and $b = \emptyset$ since $x \mathrel{\overline{\in}} Core$, thus $a \mathrel{\overline{\in}} Set$. Therefore $a \mathrel{\overline{\in}} y$, but $y \mathrel{\overline{\in}} Set$ so $a \in_{\emptyset} y$. Similarly $a \in_b y$ implies $a \in_b x$, so $x = y$ by Extensionality for multisets.

- Comprehension

  $$(\forall x \mathrel{\overline{\in}} Set)(\exists y \mathrel{\overline{\in}} Set)(\forall a \mathrel{\overline{\in}} Set)(a \mathrel{\overline{\in}} y \Leftrightarrow (a \mathrel{\overline{\in}} x \wedge \phi(a)))$$

  By Comprehension for multisets we have $y := \{a \otimes b : a \in_b x \wedge \phi(a)\}$. It is enough to show $y \mathrel{\overline{\in}} Set$, but $x \mathrel{\overline{\in}} Set$ so $a \mathrel{\overline{\in}} Set$ and $b = \emptyset$ whenever $a \in_b y$.

- Replacement

  $$(\forall x \mathrel{\overline{\in}} Set)\Big((\forall a \mathrel{\overline{\in}} x)(\exists! b \mathrel{\overline{\in}} Set)\phi(a, b) \Rightarrow (\exists y \mathrel{\overline{\in}} Set)(\forall b \mathrel{\overline{\in}} Set)(b \mathrel{\overline{\in}} y \Leftrightarrow (\exists a \mathrel{\overline{\in}} x)\phi(a, b))\Big)$$

  By Replacement for multisets we have $y := \{b \otimes \emptyset : (\exists a)(a \mathrel{\overline{\in}} x \wedge \phi(a, b))\}$, and clearly $y \mathrel{\overline{\in}} Set$.

- Pair Set

  $$(\forall x, y \mathrel{\overline{\in}} Set)(\exists a \mathrel{\overline{\in}} Set)(\forall b \mathrel{\overline{\in}} Set)(b \mathrel{\overline{\in}} a \Leftrightarrow (b = x \vee b = y))$$

  If $x, y \mathrel{\overline{\in}} Set$, then clearly $\{x \otimes \emptyset, y \otimes \emptyset\} \mathrel{\overline{\in}} Set$.

- Union

  $$(\forall x \mathrel{\overline{\in}} Set)(\exists y \mathrel{\overline{\in}} Set)(\forall a \mathrel{\overline{\in}} Set)(a \mathrel{\overline{\in}} y \Leftrightarrow (\exists b \mathrel{\overline{\in}} x) a \mathrel{\overline{\in}} b)$$

  It is enough to show that $x \mathrel{\overline{\in}} Set \Rightarrow \bigcup x \mathrel{\overline{\in}} Set$.

  Clearly $a \mathrel{\overline{\in}} Set$ for any $a \mathrel{\overline{\in}} \bigcup x$, and $\frac{\bigcup x}{a} = \emptyset$ since $(\forall y \mathrel{\overline{\in}} x)(a \mathrel{\overline{\in}} y \Rightarrow \frac{y}{a} = \emptyset)$.

- Power Set

$$(\forall x \mathrel{\overline{\in}} Set)(\exists y \mathrel{\overline{\in}} Set)(\forall a \mathrel{\overline{\in}} Set)(a \mathrel{\overline{\in}} y \Leftrightarrow a \mathrel{\overline{\subset}} x)$$

If $x \mathrel{\overline{\in}} Set$, then trivially $(\forall y \mathrel{\overline{\subset}} x) y \mathrel{\overline{\in}} Set$ so $\mathcal{P}x \mathrel{\overline{\in}} Set$.

- Infinity

Clearly $\boldsymbol{\omega} \mathrel{\overline{\in}} Set$ and corresponds to the von Neumann ω under our interpretation.

□

Remark 41. *Note that Definition 32 ensures that* $(\forall \alpha \mathrel{\overline{\in}} ON)\alpha \mathrel{\overline{\in}} Set$.

Remark 42. *Up to this point, since we have chosen our axioms to avoid all manipulation of multiplicities so far, note that any model of Zermelo-Fraenkel set theory can be trivially regarded as a model for our multiset theory by interpreting* $x \in y$ *as* $x \in_{\emptyset} y$ *(we could replace* $\emptyset$ *everywhere by another designated set if we rewrite the Axioms of Pair Set, Power Set and Infinity for multisets to avoid specifying* $\boldsymbol{\emptyset}$ *as the required multiplicity). Hence the consistency strength of our theory is the same as ZF.*

3.1.7 Multiplicity Replacement

By adding different axioms to handle multiplicities, we can extend MS to implement different systems of multisets of various strengths. For example consider the following theory:

Axiom 18. *Axiom of Finite Multiplicities*

$$(\forall x, y, a)(x \in_a y \Rightarrow a \mathrel{\overline{\in}} \boldsymbol{\omega}) \wedge (\forall x)(\forall a \mathrel{\overline{\in}} \boldsymbol{\omega})(\exists y) y = \{x \otimes a\}$$

Definition 37. *The theory* $MS_{\boldsymbol{\omega}}$ *consists of all MS axioms, plus the Axiom of Finite Multiplicities.*

$\mathrm{MS}_{\boldsymbol{\omega}}$ describes a system of multisets with finite multiplicities which is essentially equivalent to the theory MST in [Blizard 2]. The one major difference is that in MST the multiplicity 0 is regarded as non-membership while our theory allows positive membership with multiplicity 0. In other words the multiplicity n in $\mathrm{MS}_{\boldsymbol{\omega}}$ corresponds to the multiplicity $n+1$ in MST.

If we want a theory of multiset where multiplicities are ZF cardinals (similar to the theory MSTC presented in Blizard[3]), we make use of the fact that the class Set is a model of ZF. Thus we let $Card$ be the class of cardinals in Set and add to MS the axiom schema below:

Axiom Schema 19.

$$(\forall x)\Big((\forall a \mathrel{\overline{\in}} x)(\exists! b \mathrel{\overline{\in}} Card)\phi(a,b) \Rightarrow (\exists y)y = \{a \otimes b : a \mathrel{\overline{\in}} x \wedge \phi(a,b)\}\Big)$$

for any formula ϕ with two free variables and possibly parameters.

This allows us to replace multiplicities in existing multisets with any ZF cardinals, as long as the replacement is a definable map. To ensure that any multiplicity is a ZF cardinal we add the following axiom

Axiom 20. $(\forall x, y, a)(x \in_a y \Rightarrow a \mathrel{\overline{\in}} Card)$

Remark 43. *With the way ordinals are defined in our multiset theory, every multiset ordinal is automatically in the class Set. As long as our axiom regarding multiplicities is strong enough to replace every multiplicity by $\emptyset$, two ordinals will have a bijection in the universe of multisets if and only if they have a bijection in the class Set, and so the alephs in the multiset model (defined analogously to their set-theoretic counterparts) will be the same as the set-theoretic alephs in the ZF interpretation of the class Set.*

In order to set up a stratification system for multisets later on we will need the strongest possible axiom to manipulate multiplicities. Namely, for any multiset and any definable map from it to the universe of multisets we can use that map to replace the multiplicities in the given multiset at will.

Axiom Schema 21. *Axiom schema of Multiplicity Replacement*

$$(\forall x)\Big((\forall a \mathrel{\overline{\in}} x)(\exists! b)\phi(a,b) \Rightarrow (\exists y)y = \{a \otimes b : a \mathrel{\overline{\in}} x \wedge \phi(a,b)\}\Big)$$

for any formula ϕ with two free variables and possibly parameters.

The addition of this axiom means that our theory can no longer accept models of Zermelo-Fraenkel set theory since set-theoretic Extensionality is provably false. For example $\{\emptyset \otimes \emptyset\}$ and $\{\emptyset \otimes \{\emptyset \otimes \emptyset\}\}$ are proved to exist as distinct multisets with exactly the same members. In the same way, it cannot accept models of any traditional multiset theories where multiplicities are integers or cardinals, since with Multiplicity Replacement it is easy to create multiplicities that would be disallowed by those theories. However we will show that the schema of Multiplicity Replacement does not raise the consistency strength of the theory by constructing a model for it from any ZF model.

Definition 38. *The theory MS^+ consists of all MS axioms, plus the schema of Multiplicity Replacement.*

Lemma 61. *(MS$^+$)*

$$(\forall x)(\forall a \in \bigcup x)\frac{\bigcup x}{a} = \bigcup\{\tfrac{y}{a} \otimes b : y \in_b x \wedge a \mathrel{\overline{\in}} y\}$$

Proof. By Multiplicity Replacement let

$$A := \Big\{v \otimes w : v \mathrel{\overline{\in}} \bigcup x \wedge w = \bigcup\{\tfrac{y}{v} \otimes b : y \in_b x \wedge v \mathrel{\overline{\in}} y\}\Big\}$$

It is easy to show that $y \mathrel{\overline{\subset}} A$ for all $y \mathrel{\overline{\in}} x$, so $\bigcup x \mathrel{\overline{\subset}} A$. Thus $(\forall a \mathrel{\overline{\in}} \bigcup x)\frac{\bigcup x}{a} \mathrel{\overline{\subset}} \frac{A}{a}$.

Conversely if $a \mathrel{\overline{\in}} y \mathrel{\overline{\in}} x$, then $y \mathrel{\overline{\subset}} \bigcup x$ so $\frac{y}{a} \mathrel{\overline{\subset}} \frac{\bigcup x}{a}$. Hence $\frac{A}{a} \mathrel{\overline{\subset}} \frac{\bigcup x}{a}$.

The claims thus follows by antisymmetry of $\overline{\subset}$. □

3.1.8 Well-founded multisets

Start with a model of MS$^+$.

We say a multiset x is well-founded if every sub-multiset A of the $\overline{\subset}$-minimal multiset containing x (see Corollary 9) has a minimal member y such that

$$(\forall z \mathrel{\overline{\in}} y)(z \mathrel{\overline{\notin}} A \wedge \tfrac{y}{z} \mathrel{\overline{\notin}} A)$$

Alternatively, we can define the class WF of well-founded multisets by the analogue of the cumulative hierarchy:

$$\begin{aligned}
V_{\emptyset} &:= \emptyset \\
V_{\alpha+1} &:= \Big\{x \otimes \emptyset : (\forall y, a)\Big(y \in_a x \Rightarrow (y \mathrel{\overline{\in}} V_\alpha \wedge a \mathrel{\overline{\in}} V_\alpha)\Big)\Big\} \\
V_\lambda &:= \bigcup\{V_\alpha \otimes \emptyset : \alpha < \lambda\} \text{ for limit } \lambda \\
x \mathrel{\overline{\in}} WF &\Leftrightarrow_{df} (\exists \alpha \mathrel{\overline{\in}} ON) x \mathrel{\overline{\in}} V_\alpha
\end{aligned}$$

A simple induction shows that the V_α are all closed and form a nested hierarchy. Furthermore also by induction WF is precisely the class of well-founded multisets. We can thus define the rank of a well-founded multiset as the minimal ordinal α such that V_α contains the multiset in question, with the property that

$$\operatorname{rank} x := \max\{\sup_{y \mathrel{\overline{\in}} x} \operatorname{rank} y + 1, \sup_{y \mathrel{\overline{\in}} x} \operatorname{rank} \tfrac{x}{y} + 1\}$$

Lemma 62. $(\forall x)((\forall y \mathrel{\overline{\in}} x)(y \mathrel{\overline{\in}} WF \wedge \frac{x}{y} \mathrel{\overline{\in}} WF) \Rightarrow x \mathrel{\overline{\in}} WF)$

Proof. Let $\alpha := \max\{\sup_{y \mathrel{\overline{\in}} x} \operatorname{rank} y + 1, \sup_{y \mathrel{\overline{\in}} x} \operatorname{rank} \frac{x}{y} + 1\}$, then $x \mathrel{\overline{\in}} V_{\alpha+1}$. □

Lemma 63. $(\forall x \mathrel{\overline{\in}} WF)(\forall y \mathrel{\overline{\subset}} x) y \mathrel{\overline{\in}} WF$

Proof. If $x \mathrel{\overline{\in}} WF$ and $y \mathrel{\overline{\subset}} x$ but $y \mathrel{\overline{\notin}} WF$, let x be of minimal rank. Then for all $z \mathrel{\overline{\in}} y$, $z \mathrel{\overline{\in}} WF$ and $\frac{y}{z} \mathrel{\overline{\in}} WF$ since $\frac{y}{z} \mathrel{\overline{\subset}} \frac{x}{z}$. Hence $y \mathrel{\overline{\in}} WF$, contradiction. □

Lemma 64. $(\forall \alpha \mathrel{\overline{\in}} ON)\alpha \mathrel{\overline{\in}} WF$

Proof. We already showed that ordinals are closed, so it is straightforward to see that α^+ is the $\overline{\subset}$-minimal multiset that has α as a member. Suppose $A \mathrel{\overline{\subset}} \alpha^+$ and let y be the $\overline{\in}$-minimal member of A.

If there is some $z \mathrel{\overline{\in}} y$, then clearly $z \mathrel{\overline{\notin}} A$. Furthermore y is a non-empty ordinal by Lemma 56, so $\frac{y}{z} = \emptyset$. But $\emptyset \mathrel{\overline{\in}} y$ as in the proof of Lemma 56, so $\emptyset \mathrel{\overline{\notin}} A$ by minimality of y. □

Axiom 22. *(Axiom of Foundation)* $(\forall v)v \mathrel{\overline{\in}} WF$

Just like in set theory, we can show that:

Theorem 5. *WF is a model of MS$^+$ plus Foundation.*

Proof. The proof is much like its counterpart in set theory. Extensionality holds since WF is closed downwards. Comprehension, Pairing, Power Set and Multiplicity Replacement all hold by Lemma 62. Union holds by an induction on rank, using the recursive relation in Lemma 61. Replacement follows from Union and lemma 62. Infinity holds since $\boldsymbol{\omega}$ is well-founded by Lemma 64. Foundation holds since the cumulative hierarchy defined relative to WF is exactly the same as the cumulative hierarchy that forms WF. □

Corollary 11. *If MS$^+$ is consistent, then MS$^+$ plus Foundation is consistent.*

3.1.9 Transitive closed multisets

We have shown in the theory MS that for any multiset x there is a transitive multiset containing x and a closed multiset containing x. We now show in MS$^+$ that there is always a transitive closed multiset containing x.

Lemma 65. *(MS$^+$) For any multiset x there is a transitive closed multiset y such that $x \mathrel{\overline{\in}} y$.*

Proof. Let a be closed such that $\{x \otimes \emptyset\} \mathrel{\overline{\in}} a$, and let $b = \bigcup a$. Then $x \mathrel{\overline{\in}} b$.

If $v \mathrel{\overline{\in}} b$, then $v \mathrel{\overline{\in}} w$ for some $w \mathrel{\overline{\in}} a$ so $v \mathrel{\overline{\in}} a$. Thus $v \mathrel{\overline{\subset}} b$, i.e. b is transitive. Also

$$\frac{b}{v} = \bigcup \{\tfrac{w}{v} \otimes \emptyset : v \mathrel{\overline{\in}} w \mathrel{\overline{\in}} a\} \mathrel{\overline{\subset}} b$$

since $\frac{w}{v} \mathrel{\overline{\in}} a$ if $v \mathrel{\overline{\in}} w \mathrel{\overline{\in}} a$.

Let c be the smallest closed multiset such that $b \mathrel{\overline{\in}} c$. Then an easy induction on the minimal property of c shows that for all $w \mathrel{\overline{\in}} c$

$$w \mathrel{\overline{\subset}} b \wedge (\forall v \mathrel{\overline{\in}} w) \tfrac{w}{v} \mathrel{\overline{\subset}} b$$

If $w = b$ the claim is trivial.

Suppose the claim is true for w and $v \mathrel{\overline{\in}} w$. Then $v \mathrel{\overline{\in}} b$, so $v \mathrel{\overline{\subset}} b$ and thus

$$(\forall q \mathrel{\overline{\in}} v) \tfrac{v}{q} \mathrel{\overline{\subset}} \tfrac{b}{q} \mathrel{\overline{\subset}} b$$

Also by the claim $\frac{w}{v} \mathrel{\overline{\subset}} b$ so

$$(\forall q \mathrel{\overline{\in}} \tfrac{w}{v}) \frac{\frac{w}{v}}{q} \mathrel{\overline{\subset}} \tfrac{b}{q} \mathrel{\overline{\subset}} b$$

and the induction is complete.

By Multiplicity Replacement let $y := \{v \otimes b : v \mathrel{\overline{\in}} c\}$, then y is also closed and obviously $x \mathrel{\overline{\in}} y$.

If $w \mathrel{\overline{\in}} y$, then $w \mathrel{\overline{\in}} c$ so

$$(\forall v \mathrel{\overline{\in}} w) \tfrac{w}{v} \mathrel{\overline{\subset}} b = \tfrac{y}{v}$$

Hence $w \mathrel{\overline{\subset}} y$, i.e. y is transitive as required. □

Remark 44. *Any model of Zermelo set theory that refutes transitive containment (such as that given in [Mathias 1]) readily provides a model of MS minus Replacement in which some multiset is not contained in any transitive multiset. It remains to be investigated whether the lemma above still holds if the underlying theory is weakened to MS.*

3.1.10 Stratification, Coret's Lemma and hereditarily symmetric multisets

Throughout this section we work in a model of MS$^+$ plus Foundation.

In this section we extend the language of multisets $\mathcal{L}_{\mathcal{H}}$ (see Definition 22) to a new language $\mathcal{L}_{\mathcal{H}}{}^+$ by introducing $\overline{\subset}$ as a formal symbol, whose meaning is the same as before. Thus formally we add the axiom

Axiom 23.

$$x \mathrel{\overline{\subset}} y \Leftrightarrow x = y \vee (\exists R)\Big(\langle x, y\rangle \mathrel{\overline{\in}} R \wedge (\forall v, w)(\langle v, w\rangle \mathrel{\overline{\in}} R \Rightarrow (\forall a \mathrel{\overline{\in}} v)(a \mathrel{\overline{\in}} w \wedge \langle \tfrac{v}{a}, \tfrac{w}{a}\rangle \mathrel{\overline{\in}} R))\Big)$$

Remark 45. *As an alternative approach, we could have introduced $\overline{\subset}$ as a new symbol from the beginning with the associated axiom schema:*

- $\overline{\subset}$ *is a partial order*
- $(\forall x, y)(x \overline{\subset} y \Leftrightarrow (\forall a \overline{\in} x)(a \overline{\in} y \wedge \frac{x}{a} \overline{\subset} \frac{y}{a})$
- $(\forall x, y)\Big(\phi(x, y) \Leftrightarrow (\forall a \overline{\in} x)(a \overline{\in} y \wedge \phi(\frac{x}{a}, \frac{y}{a}))\Big) \Rightarrow (\forall x, y)(\phi(x, y) \Rightarrow x \overline{\subset} y)$ *for all formula ϕ with two variables.*

Then we can state the axioms of Union, Replacement, Power Set and Infinity using the $\overline{\subset}$ symbol directly instead of the previous definition for the inclusion relation. Finally we can prove that the previous definition holds for x and y exactly when $x \overline{\subset} y$, in the same way that we proved the properties of the inclusion relation.

Definition 39. *If ϕ is a formula in the extended language, a stratification σ of ϕ is a map from the set of variables in ϕ to $\boldsymbol{\omega} \times \boldsymbol{\omega}$ such that:*

- *If $x = y$ occurs in ϕ, then $\sigma(x) = \sigma(y)$.*
- *If $x \in_a y$ occurs in ϕ and $\sigma(x) = \langle m, n\rangle$, then $\sigma(y) = \langle m+1, n\rangle$ and $\sigma(a) = \langle m, n+1\rangle$.*
- *If $x \overline{\subset} y$ occurs in ϕ, then $\sigma(x) = \sigma(y)$, with the caveat that $\sigma(x) \neq \langle 0, 0\rangle$.*

We say the formula ϕ is stratified if it has a stratification σ, and say x has type $\sigma(x)$ for a variable x in ϕ.

Remark 46. *The condition that $\sigma(x) \neq \langle 0, 0\rangle$ whenever $x \overline{\subset} y$ occurs is to make sure that the proof of Coret's Lemma for multisets (see below) goes through. If $\sigma(x) = \langle 0, 0\rangle$ was allowed, then the lemma would imply $f(x) \overline{\subset} f(y)$ for any bijective function-class f, which is false.*

Definition 40. *The operator j*
Suppose f is a bijective function or function-class on the universe of multisets and $\langle m, n\rangle \in \mathbb{N} \times \mathbb{N}$. Define a function-class $j^{\langle m, n\rangle} f$ inductively as follows:

$$y = j^{\langle 0, 0\rangle} f(x) \Leftrightarrow_{df} y = f(x) \vee (y = x \wedge x \overline{\notin} \operatorname{dom} f)$$

$$y = j^{\langle 0, n+1\rangle} f(x) \Leftrightarrow_{df} y = \{v \otimes w : (\exists a)(a \in_w x \wedge v = j^{\langle 0, n\rangle} f(a))\}$$

$$\begin{aligned} y =& j^{\langle m+1, n\rangle} f(x) \Leftrightarrow_{df} \\ & y = \{v \otimes w : (\exists a, b)(a \in_b x \wedge v = j^{\langle m, n\rangle} f(a) \wedge w = j^{\langle m, n+1\rangle} f(b))\} \end{aligned}$$

The definitions of stratification and the operator j are both based on the idea that for any permutation f, jf acts not just on the members of a multiset but also on the multiplicities (since multiplicities are of the same type of object as sets). Combining the two definitions gives us the following analogue of Coret's Lemma.

Lemma 66. *Coret's Lemma for multisets*
Let σ be a stratification for the formula $\phi(x_1 \ldots x_n)$, then for any bijective function or function-class f

$$(\forall x_1 \ldots x_n)\phi(x_1 \ldots x_n) \Leftrightarrow \phi(j^{\sigma(x_1)}f(x_1) \ldots j^{\sigma(x_n)}f(x_n))$$

Proof. The proof is exactly the same as with the original Coret's Lemma, by induction on the formula ϕ. The only new case is with $\overline{\subset}$.

If $\phi \Leftrightarrow_{\mathit{df}} x = y$, the result follows immediately since j is well defined.

If $\phi \Leftrightarrow_{\mathit{df}} x \in_a y$, let $\sigma(x) = \langle m, n\rangle$. Then $\sigma(y) = \langle m+1, n\rangle$ and $\sigma(a) = \langle m, n+1\rangle$, so the result follows from the definition of $j^{\langle m+1, n\rangle}$.

If $\phi \Leftrightarrow_{\mathit{df}} x \mathrel{\overline{\subset}} y$, let $\sigma(x) = \sigma(y) = \langle m, n\rangle$. We prove the claim by induction on m.

If $m \geq 1$, then

$$j^{\sigma(x)}f(x) = \{j^{\langle m-1, n\rangle}f(a) \otimes j^{\langle m, n+1\rangle}f(\tfrac{x}{a}) : a \mathrel{\overline{\in}} x\}$$

$$j^{\sigma(y)}f(y) = \{j^{\langle m-1, n\rangle}f(a) \otimes j^{\langle m, n+1\rangle}f(\tfrac{y}{a}) : a \mathrel{\overline{\in}} y\}$$

Let $a \mathrel{\overline{\in}} x$, then $a \mathrel{\overline{\in}} y$ so $j^{\langle m-1, n\rangle}f(a) \mathrel{\overline{\in}} j^{\sigma(y)}f(y)$.

Furthermore $\frac{x}{a} \mathrel{\overline{\subset}} \frac{y}{a}$, so by the inductive hypothesis $j^{\langle m, n+1\rangle}f(\frac{x}{a}) \mathrel{\overline{\subset}} j^{\langle m, n+1\rangle}f(\frac{y}{a})$. Hence $j^{\sigma(x)}f(x) \mathrel{\overline{\subset}} j^{\sigma(y)}f(y)$ as required.

If $m = 0$, then $n \geq 1$. Hence

$$j^{\sigma(x)}f(x) = \{j^{\langle 0, n-1\rangle}f(a) \otimes \tfrac{x}{a} : a \mathrel{\overline{\in}} x\}$$

$$j^{\sigma(y)}f(y) = \{j^{\langle 0, n-1\rangle}f(a) \otimes \tfrac{y}{a} : a \mathrel{\overline{\in}} y\}$$

Let $a \mathrel{\overline{\in}} x$, then $a \mathrel{\overline{\in}} y$ so $j^{\langle 0, n-1\rangle}f(a) \mathrel{\overline{\in}} j^{\sigma(y)}f(y)$. But also $\frac{x}{a} \mathrel{\overline{\subset}} \frac{y}{a}$, so $j^{\sigma(x)}f(x) \mathrel{\overline{\subset}} j^{\sigma(y)}f(y)$ as required.

If $\phi \Leftrightarrow_{\mathit{df}} \psi \wedge \varphi$, then σ is a stratification for both ψ and φ. Applying the inductive hypothesis to ϕ and φ gives us the induction step.

If $\phi \Leftrightarrow_{\mathit{df}} \neg\psi$, then σ is a stratification for ψ so we can apply the inductive hypothesis to ψ.

If $\phi \Leftrightarrow_{df} (\exists x)\psi(x)$, then σ is a stratification for ψ. If $\phi(a_1 \ldots a_n)$ holds, there exists a such that $\psi(a, a_1 \ldots a_n)$ holds. By the inductive hypothesis $\psi(j^{\sigma}(x)f(a), j^{\sigma(x_1)}f(a_1) \ldots j^{\sigma(x_n)}f(a_n))$ holds, so $\phi(j^{\sigma(x_1)}f(a_1) \ldots j^{\sigma(x_n)}f(a_n))$ holds. $\square$

Thus we can reintroduce the notion of symmetry in the context of multisets. As before, let G be a group of permutations (i.e. bijections f where $\operatorname{dom} f = \operatorname{ran} f$) and $\mathcal{F}$ a filter on G. To avoid complications with multiplicities we assume that $\mathcal{F} \mathrel{\overline{\in}} Core$.

Definition 41. *For $m, n \mathrel{\overline{\in}} \boldsymbol{\omega}$, the $\langle m, n\rangle$-stabiliser of x in G is*

$$G_{\langle m,\, n\rangle}(x) = \left\{\sigma \otimes \boldsymbol{\emptyset} : \sigma \mathrel{\overline{\in}} G \wedge (\forall p, q \mathrel{\overline{\in}} \boldsymbol{\omega})\Big((p \geq m \wedge q \geq n) \Rightarrow j^{\langle p,\, q\rangle}\sigma(x) = x\Big)\right\}$$

The weak stabiliser of x in G is $G_{\boldsymbol{\omega}}(x) := \bigcup_{m,n \overline{\in} \boldsymbol{\omega}} G_{\langle m,\, n\rangle}(x)$.

Definition 42. *The multiset x is strongly symmetric if $G_{\langle m,\, n\rangle}(x) \mathrel{\overline{\in}} \mathcal{F}$ for some $n \mathrel{\overline{\in}} \boldsymbol{\omega}$, and weakly symmetric if $G_{\boldsymbol{\omega}}(x) \mathrel{\overline{\in}} \mathcal{F}$.*

Lemma 67. *Let $\phi(x)$ be a stratified formula with all parameters strongly (or weakly) symmetric. If $(\exists! x)\phi(x)$, then that unique witness x is strongly (respectively weakly) symmetric.*

Proof. Let $a_1 \ldots a_l$ be the parameters of $\phi(x)$, where a_i has type $\langle p_i, q_i\rangle$ and x has type $\langle p, q\rangle$ in some stratification of ϕ. Suppose $\phi(x, a_1 \ldots a_l)$ holds.

If the a_i are strongly symmetric and $G_{\langle m(i),\, n(i)\rangle}(a_i) \mathrel{\overline{\in}} \mathcal{F}$, then $H := \bigcap_i G_{\langle m(i),\, n(i)\rangle}(a_i) \mathrel{\overline{\in}} \mathcal{F}$.

Let $m := \max\{m(1) \ldots m(l)\}$ and $n := \max\{n(1) \ldots n(l)\}$.

By Coret's Lemma, for any $a, b \geq 0$ and $\sigma \mathrel{\overline{\in}} H$

$$\phi(j^{\langle p + m + a,\, q + n + b\rangle}\sigma(x), j^{\langle p_1 + m + a,\, q_1 + n + b\rangle}\sigma(a_1) \ldots j^{\langle p_l + m + a,\, q_l + n + b\rangle}\sigma(a_l))$$

But this is exactly $\phi(j^{\langle p + m + a,\, q + n + b\rangle}\sigma(x), a_1 \ldots a_l)$, so by uniqueness of x we have $x = j^{\langle p + m + a,\, q + n + b\rangle}\sigma(x)$. Hence $H \mathrel{\overline{\subset}} G_{\langle p + m,\, q + n\rangle}(x)$ and thus x is strongly symmetric.

If the a_i are weakly symmetric, take

$$H := \bigcap_i G_{\boldsymbol{\omega}}(a_i) \mathrel{\overline{\in}} \mathcal{F}$$

If $\sigma \in H$, let $\sigma \mathrel{\overline{\in}} G_{\langle m(i),\, n(i)\rangle}(a_i)$ for each a_i and $m := \max\{m(1) \ldots m(l)\}$, $n := \max\{n(1) \ldots n(l)\}$.

By the same argument as above, $x = j^{\langle p + m + a,\, q + n + b\rangle}\sigma(x)$ for any $a, b \geq 0$, so $H \mathrel{\overline{\subset}} G_{\omega}(x)$. $\square$

Let HSM be the class of hereditarily strongly symmetric multisets, i.e. $x \mathrel{\overline{\in}} HSS$ if there is a closed set y containing x such that everything in y is strongly symmetric. Similarly we define HWM to be the class of hereditarily weakly symmetric multisets. As with HSS and HWS, these are formally definable classes in the language of multisets.

Remark 47. *In the definition above we use the notion of closed set instead of transitive set since if x is hereditarily symmetric, it makes sense to require all multiplicities involved in x also to be hereditarily symmetric.*

Define the strong closure condition as the exact analogue of our previous definition

$$\{\sigma \times \emptyset : \sigma \mathrel{\overline{\in}} G \wedge (\forall m, n \mathrel{\overline{\in}} \omega)(\forall x \mathrel{\overline{\in}} HSM)(j^{\langle m, n\rangle}\sigma(x) \mathrel{\overline{\in}} HSM)\} \mathrel{\overline{\in}} \mathcal{F}$$

and define the weak closure condition accordingly.

Theorem 6. *HSM and HWM are models of the stratified axioms of MS^+ plus Foundation if the corresponding closure conditions hold.*

Proof. The proof is exactly the same as with sets, so we only give an outline.

Extensionality and Foundation are inherited from the starting model. By the lemma above it is trivial to verify that stratified Δ_0-Comprehension, Empty Set and Pairing hold.

Union follows from the same lemma since the definition of union, written in terms of $\overline{\subset}$, is stratified. In fact this is very much the reason we introduced an extra formal symbol for $\overline{\subset}$ and defined the j operator around it.

By the closure condition $\{y \times \emptyset : y \mathrel{\overline{\in}} HSM \wedge y \mathrel{\overline{\subset}} x\}$ is hereditarily symmetric and acts as the power set of x in HSM, and the same holds for HWM. Thus Power Set holds in both structures.

Also by the closure condition $V_\alpha \cap HSM \mathrel{\overline{\in}} HSM$ and $V_\alpha \cap HWM \mathrel{\overline{\in}} HWM$ for large α, so stratified Comprehension can be reduced to stratified Δ_0-Comprehension. In addition, this allows us to prove the multiset equivalent of the ZF axiom of Collection. Once we have established stratified Multiplicity Replacement, together with stratified Comprehension and Collection we can prove stratified Replacement.

Infinity (in a stratified form) holds by choosing a suitable hereditarily symmetric well-order that is externally isomorphic to $\boldsymbol{\omega}$.

The only case not present in the set-theoretic analogue of this proof is stratified Multiplicity Replacement, i.e. stratified instances of the schema

$$(\forall x)\Big((\forall a \mathrel{\overline{\in}} x)(\exists! b)\phi(a, b) \Rightarrow (\exists y) y = \{a \otimes b : a \mathrel{\overline{\in}} x \wedge \phi(a, b)\}\Big)$$

Let $\phi(a,b)$ be a stratified formula defining a function from $x \mathrel{\overline{\in}} HSM$ to HSM in which a has type $\langle m,n\rangle$ and b has type $\langle m,n+1\rangle$. Assume that ϕ has no parameters; the general case can be dealt with similarly by raising the types in the proof above the types of the parameters. We show

$$y := \{a \otimes b : a \mathrel{\overline{\in}} x \wedge \phi(a,b)\} \mathrel{\overline{\in}} HSM$$

It is enough to prove that y is symmetric. Suppose $G_{\langle p,\, q\rangle}(x) \mathrel{\overline{\in}} \mathcal{F}$. By Coret's Lemma

$$(\forall i,j \geq 0)(\forall \sigma \mathrel{\overline{\in}} G)\Big(\phi(a,b) \Leftrightarrow \phi(j^{\langle p+m+i,\, q+n+j\rangle}\sigma(a), j^{\langle p+m+i,\, q+n+j+1\rangle}\sigma(b))\Big)$$

But

$$(\forall i,j \geq 0)(\forall \sigma \mathrel{\overline{\in}} G_{\langle p,\, q\rangle}(x))(\forall a \mathrel{\overline{\in}} x) j^{\langle p+m+i,\, q+n+j\rangle}\sigma(a) \mathrel{\overline{\in}} x$$

Hence $G_{\langle p+m+1,\, q+n\rangle}(y) \supset G_{\langle p,\, q\rangle}(x) \mathrel{\overline{\in}} \mathcal{F}$ and y is strongly symmetric.

If $y := \{a \otimes b : a \mathrel{\overline{\in}} x \wedge \phi(a,b)\} \mathrel{\overline{\in}} HWM$, a similar argument shows that $G_{\boldsymbol{\omega}}(y) \supset G_{\boldsymbol{\omega}}(x)$ so y is weakly symmetric. □

3.2 A model for the theory

Start with a model V of ZF. It is possible and indeed quite natural to construct a well-founded model for our multiset theory from V by interpreting $x \in_a y$ in the language of multiset as $\langle x, a\rangle \in y$ in a recursively defined subclass of V. However we will take a different approach and construct a graph-based model in which there are interesting manifestations of anti-foundation, in the hope that this will give rise to some consistency results for NF-like multiset theories (where we know that foundation must fail).

For our purposes we will only consider directed 3-uniform hypergraphs, which we simply call hypergraphs for brevity. The model for our multiset theory will be a definable class of these hypergraphs.

We implement a hypergraph H as the set of its edges, where each edge is an ordered tuple $\langle x,y,z\rangle$. Write $H(x,y,z)$ as shorthand for $\langle x,y,z\rangle \in H$.

Definition 43. *Let H be a hypergraph and x any set.*

Let

$$H^{-1}x := \{y : (\exists z)\langle x,y,z\rangle \in H\}$$

and

$$H^{-2}x := \{z : (\exists y)\langle x,y,z\rangle \in H\}$$

Similarly let

$$H^{-1} := \{y : (\exists x,z)\langle x,y,z\rangle \in H\}$$

and

$$H^{-2} := \{z : (\exists x, y)\langle x, y, z\rangle \in H\}$$

Definition 44.

$$x_H := \bigcap\{X : x \in X \wedge (\forall y \in X)(H^{-1}y \subset X \wedge H^{-2}y \subset X)\}$$

In other words x_H contains x and all vertices of H accessible from x by a finite directed path.

Let H_x be the restriction of H to x_H. It is the smallest subgraph of H containing x and closed under outward edges.

Definition 45. *A pointed hypergraph $[H, h]$ is an ordered pair $\langle H, h\rangle$ where H is a hypergraph and h a set which we call the point of $[H, h]$. We say $[G, h] \cong [H, h]$ if there is an isomorphism between G and H that takes g to h.*

Definition 46. *If $[H, h]$ is a pointed hypergraph, let*

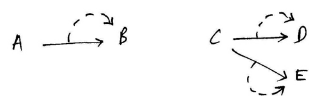
$$\begin{aligned}\mathrm{Dom}[H, h] := \{h\} &\cup \{x : (\exists y, z)\langle x, y, z\rangle \in H\}\\ &\cup \{y : (\exists x, z)\langle x, y, z\rangle \in H\}\\ &\cup \{z : (\exists x, y)\langle x, y, z\rangle \in H\}\end{aligned}$$

Definition 47. *A finite directed path in $[H, h]$ is a finite sequence $x_1 \ldots x_n$ in $\mathrm{Dom}[H, h]$ where $x_{i+1} \in H^{-1}x_i$ or $x_{i+1} \in H^{-2}x_i$. We say that the directed path is from x_1 to x_n.*

Definition 48. *$[H, h]$ is accessible if $\mathrm{Dom}[H, h] = h_H$. Equivalently, for any $x \in \mathrm{Dom}[H, h]$ there is a finite directed path in $[H, h]$ from h to x.*

Remark 48. *For any hypergraph H and any x, $[H_x, x]$ is accessible.*

To help with illustrations, we introduce a graphical representation of pointed hypergraphs. We represent the point with a star and and edge $\langle x, y, z\rangle$ by a solid arrow from x to y with a segmented arrow branching out to z. The intended interpretation is that (the object represented by) y belongs to x with multiplicity z. We omit the names of the vertices unless it is necessary to specify them.



Both graphs in the figure above depict the singleton of the empty multiset with empty multiplicity, but the second graph has two vertices both standing for the empty multiset. Obviously there will be many non-isomorphic graphs depicting the same multiset, so we need to define an equivalence relation to enforce extensionality.

Definition 49. *For any relation* $\sim \subset \mathrm{Dom}[G,g] \times \mathrm{Dom}[H,h]$*, say* $\sim$ *is a bisimulation between* $[G,g]$ *and* $[H,h]$ *if for any* $a \sim x$ *we have*

$$(\forall b, c \in \mathrm{Dom}[G,g])(G(a,b,c) \Rightarrow (\exists y, z \in \mathrm{Dom}[H,h])(H(x,y,z) \wedge b \sim y \wedge c \sim z))\wedge$$
$$(\forall y, z \in \mathrm{Dom}[H,h])(H(x,y,z) \Rightarrow (\exists b, c \in \mathrm{Dom}[G,g])(G(a,b,c) \wedge b \sim y \wedge c \sim z))$$

If $[G,g] = [H,h]$ *we say* $\sim$ *is a bisimulation on* $[H,h]$*.*

Lemma 68.

i If $\sim$ *is a bisimulation between* $[G,g]$ *and* $[H,h]$*, then the relation*

$$x \simeq y \Leftrightarrow_{df} y \sim x$$

is a bisimulation between $[H,h]$ *and* $[G,g]$*.*

ii If $\sim$ *is a bisimulation between* $[G,g]$ *and* $[H,h]$ *and* $\simeq$ *a bisimulation between* $[H,h]$ *and* $[Q,q]$*, then the relation*

$$x \approx d \Leftrightarrow_{df} (\exists a)(x \sim a \wedge a \simeq d)$$

is a bisimulation between $[G,g]$ *and* $[Q,q]$*.*

iii If $\sim$ *is a bisimulation between* $[G,g]$ *and* $[H,h]$*, then its restriction to* $\mathrm{Dom}[G_x, x] \times \mathrm{Dom}[H_y, y]$ *is a bisimulation between* $[G_x, x]$ *and* $[H_y, y]$*.*

iv Any bisimulation between $[G_x, x]$ *and* $[H_y, y]$ *is a bisimulation between* $[G,g]$ *and* $[H,h]$*.*

v Let $\sim$ *be a bisimulation between* $[G,g]$ *and* $[H,h]$ *such that* $g \sim h$*. If* $[G,g]$ *and* $[H,h]$ *are accessible, then*

$$(\forall y \in \mathrm{Dom}[H,h])(\exists x \in \mathrm{Dom}[G,g])x \sim y \wedge (\forall x \in \mathrm{Dom}[G,g])(\exists y \in \mathrm{Dom}[H,h])x \sim y$$

Remark 49. *As would be expected, since the graphs in the last figure represent the same multiset, there is a bisimulation between them, namely* $\{\langle A, C\rangle, \langle B, D\rangle, \langle B, E\rangle\}$*.*

Proof.

i The result is immediate from the definition of bisimulation.

ii Let $x \sim a \simeq d$ and $G(x,y,z)$.

There are $b, c \in \mathrm{Dom}[H,h]$ such that $y \sim b$, $z \sim c$ and $H(a,b,c)$. Hence there are $e, f \in [Q,q]$ such that $b \simeq e$, $c \simeq f$ and $Q(d,e,f)$, but then we have $y \approx e$ and $z \approx f$.

The other direction is similar, so $\approx$ is a bisimulation.

iii Let $\sim$ be a bisimulation between $[G,g]$ and $[H,h]$.

If $a \sim d$ where $a \in \mathrm{Dom}[G_x,x]$, $d \in \mathrm{Dom}[H_y,y]$ and $G_x(a,b,c)$, then clearly $G(a,b,c)$. Hence there are $e, f \in \mathrm{Dom}[H,h]$ such that $b \sim e$, $c \sim f$ and $H(d,e,f)$. But then $H_y(d,e,f)$ by definition of $[H_y,y]$.

The other direction is similar, so the restriction of $\sim$ to $\mathrm{Dom}[G_x,x] \times \mathrm{Dom}[H_y,y]$ is a bisimulation between $[G_x,x]$ and $[H_y,y]$.

iv Let $\sim$ be a bisimulation between $[G_x,x]$ and $[H_y,y]$.

If $a \sim d$ and $G(a,b,c)$, then $G_x(a,b,c)$ by definition of $[G_x,x]$. Thus there are $e, f \in \mathrm{Dom}[H_y,y]$ such that $b \sim e$, $d \sim f$ and $G_y(d,e,f)$, but then $G(d,e,f)$.

The other direction is similar, so $\sim$ is a bisimulation between $[G,g]$ and $[H,h]$.

v Let $y \in \mathrm{Dom}[H,h]$.

There exists a finite sequence $y_1 \dots y_n$ in $\mathrm{Dom}[H,h]$ where $y_1 = h$, $y_n = y$ and $y_{i+1} \in H^{-1}y_i \cup H^{-2}y_i$ for all i. We show by induction on n that there is a sequence $x_1 \dots x_n$ in $\mathrm{Dom}[G,g]$ such that $x_i \sim y_i$ and $x_{i+1} \in G^{-1}x_i \cup G^{-2}x_i$ for all i.

The case $n = 1$ is trivial.

If $n > 1$, by induction hypothesis we have a sequence $x_1 \dots x_{n-1}$ such that $x_i \sim y_i$ and $x_{i+1} \in G^{-1}x_i \cup G^{-2}x_i$ for all i. Now $y_n \in H^{-1}y_{n-1} \cup H^{-2}y_{n-1}$ and $x_{n-1} \sim yn - 1$, so there exists $x_n \in H^{-1}x_{n-1} \cup H^{-2}x_{n-1}$ such that $x_n \sim y_n$ since $\sim$ is a bisimulation.

Similarly for the other direction.

$\square$

Remark 50. *The notion of bisimulation can also be described by the following back-and-forth game:*

The game is played on a pair of accessible pointed hypergraphs $[G,g]$ and $[H,h]$. On the first turn two arbitrary vertices $G_1 \in Dom[G,g]$ and $H_1 \in Dom[H,h]$ are picked out. On the $n+1$-th turn, player 1 picks any vertex previously chosen by either player *(including the starting vertices), without loss of generality say G_m ($m \leq 2n-1$), and two more vertices G_{2n} and G_{2n+1} from the same graph, such that $\langle G_m, G_{2n}, G_{2n+1}\rangle$ is an edge of G. Player 2 then has to pick vertices H_{2n} and H_{2n+1} from the other graph, such that $\langle H_m, H_{2n}, H_{2n+1}\rangle$ is an edge of H (note that H_m has the same index as G_m). Note that both players are allowed to pick repeated vertices.*

For any pair of vertices $x \in Dom[G, g]$ and $y \in Dom[H, h]$, there is a bisimulation relating x to y if and only if player 2 has a strategy to stay alive indefinitely in the game starting with x and y: Suppose the relation $\sim$ is such a bisimulation and without loss of generality that player 1 picks an edge $\langle G_m, G_{2n}, G_{2n+1}\rangle$ where $G_m \sim H_m$. Then (assuming the Axiom of Choice) player 2 can pick an edge $\langle H_m, H_{2n}, H_{2n+1}\rangle$ where $G_{2n} \sim H_{2n}$ and $G_{2n+1} \sim H_{2n+1}$, thus keeping the situation in her favour. Conversely if player 2 has a strategy to hold out indefinitely, then it is straightforward to check the required bisimulation is the relation $G_m \sim H_m$ whenever H_m is the dictated response to G_m or G_m the response to H_m in some possible unfolding of the game.

Compared to the Ehrenfeucht-Fraïssé game (as in [Ehrenfeucht 1]), this game favours player 2 more since she does not have to keep the subgraphs on $\{G_1 \dots G_n\}$ and $\{H_1 \dots H_n\}$ isomorphic after every turn. As a result, the existence of bisimulations does not imply elementary equivalence: For example the graphs in the following diagram has a bisimulation, yet only one of them is extensional (when regarded as a relation).

Definition 50. *Say $[H, h]$ is extensional if any bisimulation on $[H, h]$ is the identity.*

Lemma 69.

i If $[H, h]$ is extensional, then so is $[H_x, x]$ for any $x \in \mathrm{Dom}[H, h]$.

ii If $[H, h]$ is extensional and $[H_x, x] \cong [H_y, y]$ for $x, y \in \mathrm{Dom}[H, h]$, then $x = y$.

Proof.

i For any $x \in \mathrm{Dom}[H, h]$ and any binary relation $\sim$ on $\mathrm{Dom}[H, h]$, $\sim$ is a bisimulation on $[H, h]$ if and only if its restriction to $x_H = \mathrm{Dom}[H_x, x]$ is a bisimulation on $[H_x, x]$ by Lemma 68. Hence if $[H, h]$ is extensional, then so is $[H_x, x]$.

ii If $[H_x, x] \cong [H_y, y]$ and ϕ is the isomorphism, then the relation $\phi(a) = b$ is a bisimulation between $[H_x, x]$ and $[H_y, y]$. By Lemma 68 it is also a bisimulation on $[H, h]$, so it is the identity and thus $x = y$.

$\square$

Lemma 70. *(Quotient Lemma) For any pointed hypergraph $[H, h]$, there exists an extensional pointed hypergraph $[Q, q]$ and a surjective quotient map $\pi : \mathrm{Dom}[H, h] \to \mathrm{Dom}[Q, q]$ such that $q = \pi(h)$,*

$$(\forall a, b, c \in \mathrm{Dom}[Q, q])\Big(Q(a, b, c) \Leftrightarrow (\exists x, y, z \in \mathrm{Dom}[H, h])(a = \pi(x) \wedge b = \pi(y) \wedge c = \pi(z))\Big)$$

and the relation $\pi(x) = y$ is a bisimulation between $[H, h]$ and $[Q, q]$. We call $[Q, q]$ the extensional quotient of $[H, h]$. Furthermore:

i The extensional quotient is unique up to isomorphism.

ii $\mathrm{Dom}[Q_\pi(x), \pi(x)] = \{\pi(y) : y \in \mathrm{Dom}[H_x, x]\}$ for any $x \in \mathrm{Dom}[H, h]$
In particular if $[H, h]$ is accessible, then so is $[Q, q]$.

iii For any $x \in \mathrm{Dom}[H, h]$, the extensional quotient of $[H_x, x]$ is (isomorphic to) $[Q_\pi(x), \pi(x)]$.

iv If $[H_x, x]$ is extensional, then $[Q_\pi(x), \pi(x)]$ is isomorphic to $[H_x, x]$.

Proof. If $\sim \subset \mathrm{Dom}[H, h]^2$, define $\sim^+$ by

$$\begin{aligned} a \sim^+ x \Leftrightarrow_{df} & (\forall b, c \in \mathrm{Dom}[H, h])(H(a, b, c) \Rightarrow (\exists y, z \in \mathrm{Dom}[H, h])(H(x, y, z) \wedge b \sim y \wedge c \sim z)) \wedge \\ & (\forall y, z \in \mathrm{Dom}[H, h])(H(x, y, z) \Rightarrow (\exists b, c \in \mathrm{Dom}[H, h])(H(a, b, c) \wedge b \sim y \wedge c \sim z)) \end{aligned}$$

Clearly $\sim_1 \subset \sim_2$ implies $\sim_1^+ \subset \sim_2^+$ and $\sim$ is a bisimulation if and only if $\sim \subset \sim^+$.

Now define $\approx \subset \mathrm{Dom}[H, h]^2$ by

$$x \approx y \Leftrightarrow_{df} (\exists \sim)(\sim \subset \sim^+ \wedge x \sim y)$$

i.e. the union of all bisimulations on $[H, h]$.

The set of relations on $\mathrm{Dom}[H, h]$ ordered by $\subset$ forms a complete lattice, and the operation $\sim \mapsto \sim^+$ is monotonic. Hence as an easy case of the Knaster-Tarski theorem we can show that $\approx$ is the same as $\approx^+$:

If $x \approx y$, then there is a relation $\sim \subset \sim^+$ such that $x \sim y$. Thus $x \sim^+ y$, so $x \approx^+ y$ by definition of $\approx^+$. Hence $\approx \subset \approx^+$ and therefore $\approx^+ \subset \approx^{++}$. Thus $\approx$ is a bisimulation, so $\approx^+ \subset \approx$ and $\approx = \approx^+$.

The identity on $\mathrm{Dom}[H, h]$ is a bisimulation, so $\approx$ is reflexive. Since $\approx$ is a bisimulation, by Lemma 68 the relations $\{\langle y, x\rangle : x \approx y\}$ and $\{\langle x, z\rangle : (\exists y) x \approx y \approx z\}$ are also bisimulations, so $\approx$ is symmetric and transitive.

Let $\mathrm{Dom}[Q,q]$ be the set of equivalence classes of $\approx$, q the equivalence class of h, and let $\pi : \mathrm{Dom}[H,h] \to \mathrm{Dom}[Q,q]$ be the corresponding quotient map. Define the relation Q on $\mathrm{Dom}[Q,q]$ by

$$Q(a,b,c) \Leftrightarrow_{df} (\exists x,y,z \in \mathrm{Dom}[H,h])(H(x,y,z) \wedge \pi(x)=a \wedge \pi(y)=b \wedge \pi(z)=c)$$

Let $\sim_Q$ be a bisimulation on $[Q,q]$. We show that $\sim_Q$ must be the identity relation.

Define a relation $\sim_H$ on $\mathrm{Dom}[H,h]$ by

$$x \sim_H y \Leftrightarrow_{df} \pi(x) \sim_Q \pi(y)$$

If $x \sim_H a$ and $H(x,y,z)$, then $Q(\pi(x),\pi(y),\pi(z))$ so there exists $d,e \in \mathrm{Dom}[Q,q]$ such that $Q(\pi(a),d,e)$ and $\pi(y) \sim_Q d$, $\pi(z) \sim_Q e$. Thus there are $b,c \in \mathrm{Dom}[H,h]$ such that $d=\pi(b)$, $e=\pi(c)$ and $H(a,b,c)$. But then $y \sim_H b$ and $z \sim_H c$.

Similarly if $x \sim_H a$ and $H(a,b,c)$, then there are $y \sim_H b$, $z \sim_H c$ such that $H(x,y,z)$. Therefore $\sim_H$ is a bisimulation.

This means $\sim_H \subset \approx$, so $x \sim_H y \Rightarrow \pi(x)=\pi(y)$ for any $x,y \in \mathrm{Dom}[H,h]$ and thus $\sim_Q$ is the identity. We have shown that $[Q,q]$ is extensional.

If $\pi(x)=a$ and $Q(a,b,c)$, there are d,e,f such that $\pi(d)=a$, $\pi(e)=b$, $\pi(f)=c$ and $H(d,e,f)$. Then $x \approx d$ so there are $y \approx e$, $z \approx f$ such that $H(x,y,z)$. But then $\pi(y)=b$ and $\pi(z)=c$.

On the other hand, if $H(x,y,z)$, then $Q(\pi(x),\pi(y),\pi(z))$. Thus the relation $\pi(x)=y$ is a bisimulation between $[H,h]$ and $[Q,q]$.

i Let $[Q,q]$ be an extensional quotient of $[H,h]$ and π the associated quotient map. Define

$$x \sim y \Leftrightarrow_{df} \pi(x)=\pi(y)$$

It suffices to show that $\sim$ is the same as the greatest bisimulation $\approx$ on $\mathrm{Dom}[H,h]$. Now

$$x \sim y \Leftrightarrow (\exists a)(\pi(x)=a \wedge a=\pi(y))$$

so $\sim$ is a bisimulation on $[H,h]$ by Lemma 68, since the relation $\pi(x)=y$ is a bisimulation. Hence $\sim \subset \approx$.

Similarly, define a relation on $\mathrm{Dom}[Q,q]$ by

$$x \simeq y \Leftrightarrow_{df} (\exists a,b)(x=\pi(a) \wedge a \approx b \wedge \pi(b)=y)$$

Then $\simeq$ is a bisimulation on $[Q,q]$ since both $\approx$ and $\pi(x)=y$ are bisimulations. But $[Q,q]$ is extensional, so $\simeq$ is the identity, and so $\approx \subset \sim$.

Thus $\sim = \approx$ as required.

ii $[H_x, x]$ and $[Q_\pi(x), \pi(x)]$ are accessible, and by Lemma 68 the relation $\pi(x) = y$ is a bisimulation between them. Hence by Lemma 68 $\mathrm{Dom}[Q_\pi(x), \pi(x)] = \pi``\mathrm{Dom}[H_x, x]$.

If $[H, h]$ is accessible, then $\mathrm{Dom}[Q_q, q] = \pi``\mathrm{Dom}[H, h] = \mathrm{Dom}[Q, q]$ so $[Q, q]$ is accessible.

iii By Lemma 68, the restriction of $\approx$ to $\mathrm{Dom}[H_x, x]^2$ is precisely the greatest bisimulation on $[H_x, x]$.

Let G be the restriction of Q to $\pi``\mathrm{Dom}[H_x, x]$, then the extensional quotient of $[H_x, x]$ as constructed in this proof is isomorphic to $[G, \pi(x)]$ via the bijection that sends the equivalence class in $\mathrm{Dom}[H_x, x]$ of any vertex a to the equivalence class of a in $\mathrm{Dom}[H, h]$.

By the result above, $[G, \pi(x)]$ is the same as $[Q_\pi(x), \pi(x)]$.

iv If $[H_x, x]$ is extensional, the restriction of $\approx$ to $\mathrm{Dom}[H_x, x] = x_H$ is the identity since it is a bisimulation by Lemma 68. Hence $\pi : \mathrm{Dom}[H_x, x] \leftrightarrow \mathrm{Dom}[Q_\pi(x), x]$ is a bijection.

Suppose $a, b, c \in \mathrm{Dom}[H_x, x]$ and $Q(\pi(a), \pi(b), \pi(c))$.

Since the relation $\pi(x) = a$ is a bisimulation, there are $d, e \in \mathrm{Dom}[H, h]$ such that $\pi(d) = \pi(b), \pi(e) = \pi(c)$ and $H(a, d, e)$. But then $d, e \in \mathrm{Dom}[H_x, x]$, so $d = b$ and $e = c$, i.e. $H(a, b, c)$.

Conversely we already know that $H(a, b, c) \Rightarrow Q(\pi(a), \pi(b), \pi(c))$, hence $[Q_\pi(x), \pi(x)] \cong [H_x, x]$.

□

Definition 51. *The hypergraph $[Q, q]$ constructed in the Quotient Lemma provides a canonical example of an extensional quotient of $[H, h]$. From now on we simply refer to it as the extensional quotient of $[H, h]$.*

Definition 52. *Say $[G, g]$ and $[H, h]$ are similar if their extensional quotients are isomorphic, and write $[G, g] \equiv [H, h]$.*

Remark 51. *It is immediate from the definition above and the Quotient Lemma that if $[G, g]$ and $[H, h]$ are extensional, then $[G, g] \equiv [H, h] \Leftrightarrow [G, g] \cong [H, h]$.*

Lemma 71. *Let $[G, g]$ and $[H, h]$ be accessible. Then $[G, g] \equiv [H, h]$ if and only if there exists a bisimulation $\sim$ between $[G, g]$ and $[H, h]$ such that $g \sim h$.*

Remark 52. *In other words $[G, g] \equiv [H, h]$ if and only if in the game described in Remark 50 with starting vertices g and h, player 2 has a strategy to hold out indefinitely.*

Proof. Let $\pi_G : [G, g] \to [P, p]$ and $\pi_H : [H, h] \to [Q, q]$ be the quotient maps.

$$\begin{tikzcd}
{[G,g]} \arrow[r, leftrightarrow, "\sim"] \arrow[d, "\pi_G"'] & {[H,h]} \arrow[d, "\pi_H"] \\
{[P,p]} \arrow[r, leftrightarrow, "\theta"] & {[Q,q]}
\end{tikzcd}$$

If θ is an isomorphism between $[P, p]$ and $[Q, q]$, then the relation $\theta(x) = y$ is clearly a bisimulation. Define

$$x \sim y \Leftrightarrow_{df} \theta\pi_G(x) = \pi_H(y)$$

Then $\sim$ is a bisimulation by Lemma 68 since the relations $\pi_G(x) = y$, $\theta(x) = y$ and $\pi_H(x) = y$ are all bisimulations. Furthermore $\theta\pi_G(g) = \theta(p) = q = \pi_H(h)$, so $g \sim h$.

Conversely, let $\sim$ be a bisimulation between $[G, g]$ and $[H, h]$ such that $g \sim h$. Define a relation $\approx \subset \mathrm{Dom}[P, p] \times \mathrm{Dom}[Q, q]$ by

$$x \approx a \Leftrightarrow_{df} (\exists y, b)(\pi_G(y) = x \wedge \pi_H(b) = a \wedge y \sim b)$$

As above $\approx$ is a bisimulation by Lemma 68.

Define a relation on $\mathrm{Dom}[P, p]$ by

$$x \simeq y \Leftrightarrow_{df} (\exists a \in \mathrm{Dom}[Q, q])(x \approx a \wedge y \approx a)$$

Then $\simeq$ is a bisimulation by Lemma 68 again. Since $[P, p]$ is extensional, $\simeq$ is the identity i.e. $\approx$ is a partial function from $\mathrm{Dom}[P, p]$ to $\mathrm{Dom}[Q, q]$.

The same reasoning with $[Q, q]$ shows that $\approx$ is injective. Moreover $p \approx q$ and the hypergraphs are accessible, so by Lemma 68 we know that $\approx$ is defined on the whole of $\mathrm{Dom}[P, p]$ and surjective on $\mathrm{Dom}[Q, q]$.

Thus $\approx$ is a bijection between $\mathrm{Dom}[P, p]$ and $\mathrm{Dom}[Q, q]$, but it is also a bisimulation between $[P, p]$ and $[Q, q]$. Furthermore $p \approx q$, so $\approx$ is an isomorphism between $[P, p]$ and $[Q, q]$. □

Corollary 12. *For any $x, y \in \mathrm{Dom}[H, h]$, $[H_x, x] \equiv [H_y, y]$ if and only if there is a bisimulation $\sim$ on $[H, h]$ such that $x \sim y$.*

Proof. Immediate from the lemma above and Lemma 68. □

Definition 53. *Say $[H, h]$ is a* multigraph *if it is accessible, extensional and*

$$(\forall a, b, c, d \in \mathrm{Dom}[H, h])(H(a, b, c) \wedge H(a, b, d) \Rightarrow c = d)$$

Write $\mathcal{H}$ for the class of multigraphs.

Lemma 72. *Let $[H, h]$ be accessible and*

$$(H(a, b, c) \wedge H(a, d, e) \wedge [H_b, b] \equiv [H_d, d]) \Rightarrow [H_c, c] \equiv [H_e, e]$$

for all $a, b, c, d, e \in \mathrm{Dom}[H, h]$. Then the extensional quotient $[Q, q]$ of $[H, h]$ is a multigraph.

Proof. Let π be the quotient map from $[H, h]$ to $[Q, q]$. By construction $[Q, q]$ is extensional, and its accessibility comes from $[H, h]$.

If $Q(\pi(a), \pi(b), \pi(c))$ and $Q(\pi(a), \pi(b), \pi(d))$, without loss of generality assume $H(a, b, c)$. Since the relation $\pi(x) = y$ is a bisimulation between $[H, h]$ and $[Q, q]$, there exist $e, f \in \mathrm{Dom}[H, h]$ such that $\pi(e) = \pi(b), \pi(f) = \pi(d)$ and $H(a, e, f)$.

The relation $\pi(x) = \pi(y)$ is the greatest bisimulation on $[H, h]$, so $[H_e, e] \equiv [H_b, b]$ and $[H_f, f] \equiv [H_d, d]$ by Corollary 12. By the hypothesis $[H_f, f] \equiv [H_c, c]$, so $[H_c, c] \equiv [H_d, d]$.

By Corollary 12 again, there is a bisimulation $\sim$ on $[H, h]$ such that $c \sim d$, so $\pi(c) = \pi(d)$ as required. □

Definition 54. *If $[H, h], [Q, q] \in \mathcal{H}$, say $[Q, q] \mathrel{\overline{\in}} [H, h]$ if there exists $d \in H^{-1}h$ such that $[Q, q] \cong [H_d, d]$.*

Definition 55. *If $[Q, q] \mathrel{\overline{\in}} [H, h]$, there is a unique $\langle h, n, v\rangle \in H$ such that $[H_n, n] \cong [Q, q]$. Let $\frac{[H,h]}{[Q,q]} := [H_v, v]$.*

Remark 53. *The definitions above overloads our symbols in the language of multisets in an obvious manner: the defined relations will interpret these symbols under our interpretation of the language.*

The language of multisets (see Definition 22) is interpreted as follows: Given any formula ϕ, form the formula $\phi^{\mathcal{H}}$ by

- Restricting all universal and existential quantifiers to the class $\mathcal{H}$.
- Replacing the identity relation with the bisimilarity relation $\equiv$.

- Replacing the membership relation $x \in_a y$ with $x \mathrel{\overline{\in}} y \wedge a = \frac{y}{x}$.

The third clause in the interpretation makes sure that the relations $y \mathrel{\overline{\in}} x$ and $a = \frac{x}{y}$ on $\mathcal{H}$ are related to the membership relation $x \in_a y$ in the same way that these symbols are defined at the beginning of Section 3.1.

Definition 56. *If ϕ is any formula in the language of multisets, write $\mathcal{H} \models \phi$ for the formula $\phi^{\mathcal{H}}$.*

We now proceed to prove all axioms of MS$^+$ under the given interpretation.

Lemma 73. *Let $[G, g], [H, h] \in \mathcal{H}$. Suppose*

$$(\forall \langle g, x, y\rangle \in G)(\exists \langle h, a, b\rangle \in H)([G_x, x] \cong [H_a, a] \wedge [G_y, y] \cong [H_b, b])$$

and

$$(\forall \langle h, a, b\rangle \in H)(\exists \langle g, x, y\rangle \in G)([G_x, x] \cong [H_a, a] \wedge [G_y, y] \cong [H_b, b])$$

Then $[G, g] \cong [H, h]$.

Proof. Define a relation $\sim \subset \mathrm{Dom}[G, g] \times \mathrm{Dom}[H, h]$ by

$$x \sim y \Leftrightarrow_{df} (x = g \wedge y = h) \vee [G_x, x] \cong [H_y, y]$$

We prove that $\sim$ is a bisimulation between $[G, g]$ and $[H, h]$. Suppose $x \sim a$ and $G(x, y, z)$.

If $x = g$ and $a = h$, then there are $b, c \in \mathrm{Dom}[H, h]$ such that $H(a, b, c)$ and $[H_b, b] \cong [G_y, y]$, $[H_c, c] \cong [G_z, z]$. But then $y \sim b$ and $z \sim c$ as required.

If $[G_x, x] \cong [H_a, a]$, let ϕ be the isomorphism. Then for $b = \phi(y)$, $c = \phi(z)$ we have $y \sim b$, $z \sim c$ and $H(a, b, c)$ as required.

The other direction is similar, so $\sim$ is a bisimulation. Since $[G, g]$, $[H, h]$ are accessible and $g \sim h$, by Lemma 71 $[G, g] \equiv [H, h]$, but they are extensional so $[G, g] \cong [H, h]$. $\square$

Lemma 74. *Axiom of Extensionality*
Let $[G, g], [H, h] \in \mathcal{H}$. Suppose

$$(\forall [Q, q] \in \mathcal{H})\,([Q, q] \mathrel{\overline{\in}} [G, g] \Leftrightarrow [Q, q] \mathrel{\overline{\in}} [H, h])$$

and

$$(\forall [Q, q] \mathrel{\overline{\in}} [H, h])\,\tfrac{[G,g]}{[Q,q]} \cong \tfrac{[H,h]}{[Q,q]}$$

Then $[G, g] \cong [H, h]$.

Proof. This is an immediate corollary of Lemma 73. $\square$

Definition 57. *For any* $[H, h] \in \mathcal{H}$ *and* $y \in H^{-1}x$*, write* $H(x, y)$ *for the unique* z *such that* $H(x, y, z)$*.*

Definition 58. *Let* $[G, g], [H, h] \in \mathcal{H}$*.*

Say $[G, g] \mathrel{\overline{\sqsubset}} [H, h]$ *if there is a relation* $\lhd \subset \mathrm{Dom}[G, g] \times \mathrm{Dom}[H, h]$ *such that for any* $x \in \mathrm{Dom}[G, g]$*,* $y \in \mathrm{Dom}[H, h]$

$$x \lhd y \Rightarrow (\forall a \in G^{-1}x)(\exists b \in H^{-1}y)([G_a, a] \cong [H_b, b] \wedge G(x, a) \lhd H(y, b))$$

and $g \lhd h$*.*

Lemma 75. *The relation* $\overline{\sqsubset}$ *respects isomorphism of multigraphs, i.e.*

$$[Q, q] \cong [G, g] \wedge [G, g] \mathrel{\overline{\sqsubset}} [H, h] \Rightarrow [Q, q] \mathrel{\overline{\sqsubset}} [H, h]$$

and

$$[Q, q] \mathrel{\overline{\sqsubset}} [G, g] \wedge [H, h] \cong [G, g] \Rightarrow [Q, q] \mathrel{\overline{\sqsubset}} [H, h]$$

Proof. If $[Q, q] \cong [G, g]$ by the isomorphism ϕ and $[G, g] \mathrel{\overline{\sqsubset}} [H, g]$ as witnessed by the relation $\lhd$, then the relation $\phi(x) \lhd y$ witnesses $[Q, q] \mathrel{\overline{\sqsubset}} [H, h]$.

If $[Q, q] \mathrel{\overline{\sqsubset}} [G, g]$ as witnessed by the relation $\lhd$ and $[H, h] \cong [G, g]$ by the isomorphism ϕ, then the relation $x \lhd \phi(y)$ witnesses $[Q, q] \mathrel{\overline{\sqsubset}} [H, h]$. $\square$

Lemma 76. *Let* $\phi(x, y)$ *be a formula in the language of multisets with two free variables and possibly parameters, such that*

$$\mathcal{H} \models (\forall x, y)(\phi(x, y) \Leftrightarrow (\forall a \mathrel{\overline{\in}} x)(a \mathrel{\overline{\in}} y \wedge \phi(\tfrac{x}{a}, \tfrac{y}{a}))$$

Then

$$(\forall [Q, q], [H, h] \in \mathcal{H})\Big(\Big(\mathcal{H} \models \phi([Q, q], [H, h])\Big) \Rightarrow [Q, q] \mathrel{\overline{\sqsubset}} [H, h]\Big)$$

Furthermore $[G, g] \mathrel{\overline{\sqsubset}} [H, h]$ *if and only if*

$$(\forall [Q, q] \mathrel{\overline{\in}} [G, g])\left([Q, q] \mathrel{\overline{\in}} [H, h] \wedge \tfrac{[G,g]}{[Q,q]} \mathrel{\overline{\sqsubset}} \tfrac{[H,h]}{[Q,q]}\right)$$

Finally $\overline{\sqsubset}$ *is reflexive, transitive and*

$$[G, g] \mathrel{\overline{\sqsubset}} [H, h] \wedge [H, h] \mathrel{\overline{\sqsubset}} [G, g] \Rightarrow [G, g] \cong [H, h]$$

i.e. $\overline{\sqsubset}$ *is antisymmetric if we interpret* $\cong$ *as the identity relation.*

Proof. For the first part, suppose $\mathcal{H} \models \phi([Q,q],[H,h])$. Define a set relation $\lhd \subset \mathrm{Dom}[Q,q] \times \mathrm{Dom}[H,h]$ by

$$x \lhd y \Leftrightarrow_{df} \mathcal{H} \models \phi([Q_x,x],[H_y,y])$$

Then it is easy to check that $\lhd$ satisfies the condition in Definition 58, so $[Q,q] \mathrel{\overline{\subset}} [H,h]$.

If $\lhd \subset \mathrm{Dom}[G,g] \times \mathrm{Dom}[H,h]$, define $\lhd^+$ by

$$x \lhd^+ y \Leftrightarrow_{df} (\forall a \in G^{-1}x)(\exists b \in H^{-1}y)([G_a,a] \cong [H_b,b] \wedge G(x,a) \lhd H(y,b))$$

Note that the operation taking $\lhd$ to $\lhd^+$ depends on the multigraphs $[G,g]$ and $[H,h]$, but to keep the notation simple we omit the associated multigraphs unless confusion may arise. It is straightforward to check that $\lhd_1 \subset \lhd_2 \Rightarrow \lhd_1^+ \subset \lhd_2^+$.

Given $[G,g]$ and $[H,h]$, let

$$x \prec y \Leftrightarrow_{df} (\exists \lhd \in \mathrm{Dom}[G,g] \times \mathrm{Dom}[H,h])(\lhd \subset \lhd^+ \wedge x \lhd y)$$

Clearly $[G,g] \mathrel{\overline{\subset}} [H,h]$ if and only if $g \prec h$.

If $\lhd \subset \lhd^+$ and $x \lhd y$, then

$$(\forall a \in G^{-1}x)(\exists b \in H^{-1}y)([G_a,a] \cong [H_b,b] \wedge G(x,a) \lhd H(y,b))$$

But then $G(x,a) \prec H(y,b)$, so $x \prec^+ y$.

This shows $\prec \subset \prec^+$, so $\prec^+ \subset \prec^{++}$ and thus $\prec = \prec^+$ by definition. Therefore

$$x \prec y \Leftrightarrow (\forall a \in G^{-1}x)(\exists b \in H^{-1}y)([G_a,a] \cong [H_b,b] \wedge G(x,a) \prec H(y,b))$$

We call $\prec$ the greatest subset relation between $[G,g]$ and $[H,h]$.

If $x \in \mathrm{Dom}[G,g]$, $y \in \mathrm{Dom}[H,h]$ and $\sim \subset \mathrm{Dom}[G_x,x] \times \mathrm{Dom}[H_y,y]$, clearly $\sim^+$ is the same relation whether defined relative to $[G,g]$ and $[H,h]$ or $[G_x,x]$ and $[H_y,y]$.

This means the restriction of $\prec$ to $\mathrm{Dom}[G_x,x] \times \mathrm{Dom}[H_y,y]$ is precisely the greatest subset relation between $[G_x,x]$ and $[H_y,y]$, hence

$$[G_x,x] \mathrel{\overline{\subset}} [H_y,y] \Leftrightarrow x \prec y$$

Now suppose for any $[Q,q] \mathrel{\overline{\in}} [G,g]$ we have $[Q,q] \mathrel{\overline{\in}} [H,h]$ and $\frac{[G,g]}{[Q,q]} \mathrel{\overline{\subset}} \frac{[H,h]}{[Q,q]}$. Let

$$x \preceq y \Leftrightarrow_{df} x \prec y \vee (x = g \wedge y = h)$$

Suppose $x \preceq y$ and $a \in G^{-1}x$.

If $x \prec y$, there exists $b \in H^{-1}y$ such that

$$[G_a, a] \cong [H_b, b] \wedge G(x, a) \prec H(y, b)$$

In particular $G(x, a) \preceq H(y, b)$.

If $x = g$ and $y = h$, then $[G_a, a] \mathrel{\overline{\in}} [G, g]$ so by the hypothesis above

$$[G_a, a] \mathrel{\overline{\in}} [H, h] \wedge \tfrac{[G,g]}{[G_a,a]} \mathrel{\overline{\sqsubset}} \tfrac{[H,h]}{[G_a,a]}$$

In other words there exists $b \in H^{-1}h$ such that

$$[G_a, a] \cong [H_b, b] \wedge [G_{G(g,a)}, G(g, a)] \mathrel{\overline{\sqsubset}} [H_{H(h,b)}, H(h, b)]$$

This shows $G(g, a) \prec H(h, b)$, so $G(g, a) \preceq H(h, b)$.

Therefore $\preceq \subset \preceq^{+}$, so $\preceq = \prec$ and thus $g \prec h$.

Conversely let $g \prec h$ and $[Q, q] \mathrel{\overline{\in}} [G, g]$, then $[Q, q] \cong [G_x, x]$ for some $x \in G^{-1}g$.

There exists $y \in H^{-1}h$ such that

$$[H_y, y] \cong [G_x, x] \wedge G(g, x) \prec H(h, y)$$

But then $[H_y, y] \cong [Q, q]$ so $[Q, q] \mathrel{\overline{\in}} [H, h]$ and

$$\tfrac{[G,g]}{[Q,q]} \cong [G_{G(g,x)}, G(g, x)] \mathrel{\overline{\sqsubset}} [H_{H(h,y)}, H(H, y)] \cong \tfrac{[H,h]}{[Q,q]}$$

so $\frac{[G,g]}{[Q,q]} \mathrel{\overline{\sqsubset}} \frac{[H,h]}{[Q,q]}$ since $\overline{\sqsubset}$ respects isomorphism.

We have shown that

$$[G, g] \mathrel{\overline{\sqsubset}} [H, h] \Leftrightarrow g \prec h \Leftrightarrow (\forall [Q, q] \mathrel{\overline{\in}} [G, g]) \Big([Q, q] \mathrel{\overline{\in}} [H, h] \wedge \tfrac{[G,g]}{[Q,q]} \mathrel{\overline{\sqsubset}} \tfrac{[H,h]}{[Q,q]}\Big)$$

If $\lhd$ is the identity, then trivially $\lhd \subset \lhd^{+}$, so $\overline{\sqsubset}$ is reflexive.

Let $\lhd_1$ witness $[Q, q] \mathrel{\overline{\sqsubset}} [G, g]$ and $\lhd_2$ witness $[G, g] \mathrel{\overline{\sqsubset}} [H, h]$. Define

$$x \lhd y \Leftrightarrow_{df} (\exists d)(x \lhd_1 d \wedge d \lhd_2 y)$$

If $x \lhd_1 d \lhd_2 y$ and $a \in G^{-1}x$, there exists $c \in G^{-1}d$ such that

$$[Q_a, a] \cong [G_c, c] \wedge Q(x, a) \lhd_1 G(d, c)$$

and there exist $b \in H^{-1}y$ such that

$$[G_c, c] \cong [H_b, b] \wedge G(d, c) \lhd_2 H(y, b)$$

But then $[Q_a, a] \cong [H_b, b]$ and $Q(x,a) \lhd H(y,b)$ so $\lhd \subset \lhd^+$.

Furthermore $q \lhd_1 g \lhd_2 h$, so $q \lhd h$ and thus $\lhd$ witnesses $[Q,q] \mathrel{\overline{\subset}} [H,h]$. This means $\overline{\subset}$ is transitive.

Let $\prec_1$ be the greatest subset relation between $[G,g]$ and $[H,h]$, and $\prec_2$ between $[H,h]$ and $[G,g]$. Define $\sim \subset \mathrm{Dom}[G,g] \times \mathrm{Dom}[H,h]$ by

$$x \sim y \Leftrightarrow_{df} x \prec_1 y \wedge y \prec_2 x$$

If $x \sim y$ and $G(x,a,b)$, there exists $c \in H^{-1}y$ such that

$$[G_a, a] \cong [H_c, c] \wedge b \prec_1 H(y,c)$$

Since $\overline{\subset}$ is reflexive and respects isomorphism

$$[G_a, a] \mathrel{\overline{\subset}} [H_c, c] \wedge [H_c, c] \mathrel{\overline{\subset}} [G_a, a]$$

Hence $a \prec_1 c \wedge c \prec_2 a$, i.e. $a \sim c$.

Let $d := H(y,c)$, then there exists $e \in G^{-1}x$ such that

$$[G_e, e] \cong [H_c, c] \wedge d \prec_2 G(x,e)$$

Then $[G_e, e] \cong [G_a, a]$ so $e = a$ and thus $b \sim d$.

Similarly if $x \sim y$ and $H(y,c,d)$, then there are $a \sim c$, $b \sim d$ such that $G(x,a,b)$. Hence $\sim$ is a bisimulation and $[G,g] \cong [H,h]$. □

This shows $\overline{\subset}$ is precisely the internal inclusion relation of $\mathcal{H}$ as defined in Definition 26 and has all of the properties specified in Axiom 12.

Lemma 77. *Axiom of Union*
Let $[H,h] \in \mathcal{H}$. There exists $[G,g] \in \mathcal{H}$ such that

$$(\forall [Q,q] \mathrel{\overline{\in}} [H,h])[Q,q] \mathrel{\overline{\subset}} [G,g]$$

and for any $[P,p] \in \mathcal{H}$

$$(\forall [Q,q] \mathrel{\overline{\in}} [H,h])[Q,q] \mathrel{\overline{\subset}} [P,p] \Rightarrow [G,g] \mathrel{\overline{\subset}} [P,p]$$

Proof. Given multigraphs $[G,g]$ and $[H,h]$ we will define a recursive relation between vertices of $[G,g]$ and sets of vertices of $[H,h]$. Intuitively the vertex a is related to the set X if the subgraph $[G_a, a]$ is the $\overline{\subset}$-least upper bound of those $[H_x, x]$ for $x \in X$. Hence $[G,g]$ is the union of $[H,h]$ if and only if g is related to $H^{-1}h$.

For any relation $\lhd \subset \mathrm{Dom}[G,g] \times \mathcal{P}(\mathrm{Dom}[H,h])$ define (cf. Lemma 61)

$$\begin{aligned} a \lhd^+ X \Leftrightarrow_{df} & (\forall [Q,q] \in \mathcal{H})([Q,q] \mathrel{\overline{\in}} [G_a,a] \Leftrightarrow (\exists x \in X)[Q,q] \mathrel{\overline{\in}} [H_x,x]) \wedge \\ & (\forall b \in G^{-1}a) G(a,b) \lhd \{H(x,y) : x \in X \wedge y \in H^{-1}x \wedge [G_b,b] \cong [H_y,y]\} \end{aligned}$$

Then it is straightforward to see that $\lhd_1 \subset \lhd_2 \Rightarrow \lhd_1^+ \subset \lhd_2^+$. Define

$$a \prec X \Leftrightarrow_{df} (\exists \lhd \subset \mathrm{Dom}[G,g] \times \mathcal{P}(\mathrm{Dom}[H,h]))(\lhd \subset \lhd^+ \wedge a \lhd X)$$

As before, by the Knaster-Tarski theorem we have $\prec = \prec^+$.

For $a, b \in \mathrm{Dom}[G,g]$ define

$$a \sim b \Leftrightarrow_{df} (\exists X \subset \mathrm{Dom}[H,h])(a \prec X \wedge b \prec X)$$

If $a \sim b$, then

$$(\forall [Q,q] \in \mathcal{H})([Q,q] \mathrel{\overline{\in}} [G_a,a] \Leftrightarrow [Q,q] \mathrel{\overline{\in}} [G_b,b])$$

But $[G,g]$ is extensional so $G^{-1}a = G^{-1}b$. Furthermore, for any $c \in G^{-1}a$ we have

$$G(a,c) \prec \{H(x,y) : x \in X \wedge y \in H^{-1}x \wedge [G_c,c] \cong [H_y,y]\}$$

and

$$G(b,c) \prec \{H(x,y) : x \in X \wedge y \in H^{-1}x \wedge [G_c,c] \cong [H_y,y]\}$$

so $G(a,c) \sim G(b,c)$.

Hence $\sim$ is a bisimulation on $[G,g]$, so it is the identity. Thus for each $X \subset \mathrm{Dom}[H,h]$, if $a \prec X$, then a is unique.

We show that if $g \prec H^{-1}h$, then $[G,g]$ is the supposed union of $[H,h]$.

Let $[Q,q] \mathrel{\overline{\in}} [H,h]$, then $[Q,q] \cong [H_d,d]$ for some $d \in H^{-1}h$.

Define $\lhd \subset \mathrm{Dom}[H,h] \times \mathrm{Dom}[G,g]$ by

$$x \lhd y \Leftrightarrow_{df} (\exists X \subset \mathrm{Dom}[H,h])(x \in X \wedge y \prec X)$$

If $x \lhd y$, then

$$(\forall [Q,q] \mathrel{\overline{\in}} [H_x,x])[Q,q] \mathrel{\overline{\in}} [G_y,y]$$

so for any $a \in H^{-1}x$ there exists $b \in G^{-1}y$ such that $[H_a,a] \cong [G_b,b]$.

But if X witnesses $x \lhd y$, then

$$H(x,a) \in \{H(x,y) : x \in X \wedge y \in H^{-1}x \wedge [G_b,b] \cong [H_y,y]\}$$

and

$$G(y,b) \prec \{H(x,y) : x \in X \wedge y \in H^{-1}x \wedge [G_b,b] \cong [H_y,y]\}$$

so $H(x,a) \lhd G(y,b)$.

Trivially $d \lhd g$, so $\lhd$ witnesses $[H_d, d] \mathrel{\overline{\subset}} [G, g]$. Therefore $[Q, q] \mathrel{\overline{\subset}} [G, g]$.

Now suppose $[Q, q] \mathrel{\overline{\subset}} [P, p]$ for any $[Q, q] \mathrel{\overline{\in}} [H, h]$.

Define $\lhd \subset \mathrm{Dom}[G, g] \times \mathrm{Dom}[P, p]$ by

$$a \lhd b \Leftrightarrow_{df} (\exists X \subset \mathrm{Dom}[H, h])\Big(a \prec X \wedge (\forall x \in X)[H_x, x] \mathrel{\overline{\subset}} [P_b, b]\Big)$$

Then $g \lhd p$ since

$$g \prec H^{-1}h \wedge (\forall x \in H^{-1}h)[H_x, x] \mathrel{\overline{\subset}} [P, p]$$

If X witnesses $a \lhd b$ and $c \in G^{-1}a$, there exists $x \in X$ and $y \in H^{-1}x$ such that $[G_c, c] \mathrel{\overline{\in}} [H_y, y]$. Let

$$Y := \{H(d, e) : d \in X \wedge e \in H^{-1}d \wedge [G_c, c] \cong [H_e, e]\}$$

then $G(a, c) \prec Y$ by the recursive property of $\prec$.

Since $[H_x, x] \mathrel{\overline{\subset}} [P_b, b]$ there exists $v \in P^{-1}b$ such that $[H_y, y] \cong [P_v, v]$, so $[G_c, c] \cong [P_v, v]$.

Suppose $d \in X$, $e \in H^{-1}d$ and $[G_c, c] \cong [H_e, e]$, then $[H_e, e] \cong [P_v, v]$.

But $[H_d, d] \mathrel{\overline{\subset}} [P_b, b]$, so $[H_{H(d,e)}, H(d, e)] \mathrel{\overline{\subset}} [P_{P(b,v)}, P(b, v)]$. Thus

$$(\forall w \in Y)[H_w, w] \mathrel{\overline{\subset}} [P_{P(b,v)}, P(b, v)]$$

Therefore $G(a, c) \lhd P(b, v)$ and so $\lhd$ witnesses $[G, g] \mathrel{\overline{\subset}} [P, p]$. We have shown that

$$g \prec H^{-1}h \Leftrightarrow [G, g] = \bigcup [H, h]$$

Now given $[H, h]$ we construct its union by adding new vertices $\nu(X)$ for all $X \subset \mathrm{Dom}[H, h]$ and build a new graph D recursively so that the extensional quotient of $[D_{\nu}(X), \nu(X)]$ is the $\overline{\subset}$-least upper bound of those $[H_x, x]$ for $x \in X$. Then we let $d := \nu(H^{-1}h)$ and the extensional quotient of $[D, d]$ will be the union of $[H, h]$. The details are as follows:

Let $\nu : \mathcal{P}\,\mathrm{Dom}[H, h] \to A$ be a bijection such that $A \cap \mathrm{Dom}[H, h] = \emptyset$ and let $d := \nu(H^{-1}h)$. A will be the set of new vertices corresponding to subsets of $\mathrm{Dom}[H, h]$, and we will build a graph C in which each vertex in A act as the union of its associated subset of $\mathrm{Dom}[H, h]$.

Let C be the smallest set such that

$$(\forall b \in \bigcup\{H^{-1}a : a \in H^{-1}h\})\langle d, b, \nu\{H(a, b) : a \in H^{-1}h \text{ such that } b \in H^{-1}a\}\rangle \in C$$

and

$$(\forall b \in \bigcup\{H^{-1}a : a \in X\})\langle \nu(X), b, \nu\{H(a,b) : a \in X \text{ such that } b \in H^{-1}a\}\rangle \in C$$

for any $X \subset \mathrm{Dom}[H,h]$ such that $\nu(X) \in \mathrm{Dom}\, C$.

Since $A \times \mathrm{Dom}[H,h] \times A$ has the same closure properties as required of C, we can form C by taking the appropriate intersection.

Informally C is built by following the recursive property of union on multiplicities (cf. Lemma 61) from the top vertex d down. Note that in the definition of C above, the vertex b does not have any descendant just yet. Intuitively b represents members of the union, their members, members of their members and so on; so now we can just copy the corresponding subgraphs of $[H,h]$ over to form the final hypergraph.

Finally the required graph D is formed by adding to C subgraphs of H so that for each vertex b that represents a member of some $x \in A$ we have $[C_b, b] = [H_b, b]$. Thus let

$$D := C \cup \bigcup\{H_b : (\exists x, y \in A) C(x, b, y)\}$$

Note that if $C(x,b,y)$, then there is no v, w such that $C(b, v, w)$. Hence $[D_b, b] = [H_b, b]$.

Now we show that $[D, d]$ has the required property.

Let $[G, g]$ be the extensional quotient of $[D, d]$ and π the quotient map. We first show $[G, g] \in \mathcal{H}$.

Clearly C_d has all the required closure properties in the definition of C, so $C_d = C$ and thus $[C, d]$ is accessible.

If $a \in \mathrm{Dom}\, H_b$ where $(\exists x, y \in A)\langle x, b, y\rangle \in C$, then a is in a finite directed path from b since $[H_b, b]$ is accessible. But b is in a finite directed path from d, so $[D, d]$ is accessible.

Now suppose $D(x, a, v)$ and $D(x, b, w)$ where $[D_a, a] \equiv [D_b, b]$.

By construction $[D_a, a] = [H_a, a]$ and $[D_b, b] = [H_b, b]$, but $[H, h]$ is extensional so $a = b$. By construction of $[D, d]$ there is a unique v such that $D(x, a, v)$ so $v = w$. Hence $[D, d]$ satisfies the conditions of Lemma 72 and $[G, g] \in \mathcal{H}$.

Finally we show that $g \prec H^{-1}h$ where $\prec$ is as defined earlier in the proof.

Define $\lhd \subset \mathrm{Dom}[G, g] \times \mathcal{P}(\mathrm{Dom}[H, h])$ by

$$a \lhd X \Leftrightarrow_{df} a = \pi\nu(X)$$

Then $g \lhd H^{-1}h$.

If $a = \pi\nu(X)$ and $v \in H^{-1}x$ for some $x \in X$, then

$$D(\nu(X), v, \nu\{H(x, y) : x \in X \text{ such that } v \in H^{-1}x\})$$

so $\pi(v) \in G^{-1}a$.

But $[D_v, v] = [H_v, v]$ is extensional, so $[G_{\pi(v)}, \pi(v)] \cong [H_v, v]$ by the Quotient Lemma.

Conversely if $b \in G^{-1}a$, then $b = \pi(v)$ where $v \in H^{-1}x$ for some $x \in X$. But then $[D_v, v] = [H_v, v]$ is extensional, so $[G_b, b] \cong [H_v, v]$.

We have shown that for any $Q \in \mathcal{H}$

$$[Q, q] \mathrel{\overline{\in}} [G_a, a] \Leftrightarrow (\exists x \in X)[Q, q] \mathrel{\overline{\in}} [H_x, x]$$

Furthermore with $b = \pi(v)$ as above, by construction

$$G(a, b) = \pi\nu(\{H(x, v) : x \in X \text{ such that } v \in H^{-1}x\})$$

But v is the unique y such that $[H_y, y] \cong [G_b, b]$ since $[H, h]$ is extensional, so

$$\{H(x, v) : x \in X \text{ such that } v \in H^{-1}x\} = \{H(x, y) : x \in X \wedge y \in H^{-1}x \wedge [G_b, b] \cong [H_y, y]\}$$

Hence $\lhd \subset \prec$ by definition, so $g \prec H^{-1}h$ as required. □

The following lemma shows that $\mathcal{H}$ is in a sense supertransitive, i.e. for any set $X \subset \mathcal{H}$ there is a multigraph whose members in the sense of $\mathcal{H}$ are precisely members of X in V.

Lemma 78. *Let ϕ be a function (a set in the ZF model V) such that* $\operatorname{dom}\phi \subset \mathcal{H}$, *$ran\phi \subset \mathcal{H}$ and*

$$(\forall [G, g], [H, h] \in \operatorname{dom}\phi)[G, g] \cong [H, h] \Rightarrow \phi[G, g] = \phi[H, h]$$

Then there exists $[Q, q] \in \mathcal{H}$ such that

$$(\forall [H, h] \in \mathcal{H})\Big([H, h] \mathrel{\overline{\in}} [Q, q] \Leftrightarrow (\exists [G, g] \in \operatorname{dom}\phi)[G, g] \cong [H, h]\Big)$$

and

$$(\forall [H, h] \in \operatorname{dom}\phi)\tfrac{[Q,q]}{[H,h]} \cong \phi[H, h]$$

Proof. If $[H, h] \in \operatorname{dom}\phi \cup \operatorname{ran}\phi$, define $\nu[H, h] \in \mathcal{H}$ as follows:

For any $a \in \operatorname{Dom}[H, h]$ let $\hat{a} := \langle a, [H, h]\rangle$, and let

$$\hat{H} := \{\langle \hat{a}, \hat{b}, \hat{c}\rangle : \langle a, b, c\rangle \in H\}$$

Let $\nu[H, h] := [\hat{H}, \hat{h}]$, then clearly $\nu[H, h] \cong [H, h]$.

Furthermore, if $[G, g]$ and $[H, h]$ are different multigraphs, then $\operatorname{dom}\nu[G, g]$ and $\operatorname{dom}\nu[H, h]$ are disjoint. Thus we can assume without loss of generality that distinct multigraphs in $\operatorname{dom}\phi \cup \operatorname{ran}\phi$ have disjoint domains.

Let $d \notin \mathrm{Dom}[G,g]$ for all $[G,g] \in \operatorname{dom}\phi \cup \operatorname{ran}\phi$. Let

$$D := \bigcup\{G : [G,g] \in \operatorname{dom}\phi \cup \operatorname{ran}\phi\} \cup \{\langle d, g, h\rangle : (\exists G, H)[H,h] = \phi[G,g]\}$$

and let $[Q,q]$ be the extensional quotient of $[D,d]$ with quotient map π.

It is straightforward to see that $[D,d]$ is accessible.

Let $D(x,a,v)$ and $D(x,b,w)$ where $[D_a,a] \equiv [D_b,b]$.

If $x = d$, by construction $[D_a,a],[D_b,b] \in \operatorname{dom}\phi$ so $[D_a,a] \cong [D_b,b]$ since they are extensional. Hence $\phi[D_a,a] = \phi[D_b,b]$ and so $v = w$.

If $x \neq d$, then $[D_x,x] \in \operatorname{dom}\phi \cup \operatorname{ran}\phi$ so trivially $a = b$ and $v = w$.

Thus $[Q,q]$ is a multigraph.

If $[H,h] \mathrel{\overline{\in}} [Q,q]$, then $[H,h] \cong [Q_\pi(a), \pi(a)]$ for some $[A,a] \in \operatorname{dom}\phi$.

Since members of $\operatorname{dom}\phi \cup \operatorname{ran}\phi$ have disjoint domains, $[D_a,a] = [A,a]$ by construction. But $[A,a]$ is extensional, so $[Q_\pi(a), \pi(a)] \cong [A,a]$ and thus $[H,h] \cong [A,a]$.

Conversely let $[H,h] \in \operatorname{dom}\phi$, then $[D_h,h] = [H,h]$ and thus $[Q_\pi(h), \pi(h)] \cong [H,h]$ since $[H,h]$ is extensional. Thus

$$[H,h] \mathrel{\overline{\in}} [Q,q] \wedge \tfrac{[Q,q]}{[H,h]} = [Q_\pi(a), \pi(a)]$$

where $[A,a] := \phi([H,h])$.

But also $[D_a,a] = [A,a]$ and $[Q_\pi(a), \pi(a)] \cong [A,a]$, so $\frac{[Q,q]}{[H,h]} \cong \phi[H,h]$. $\square$

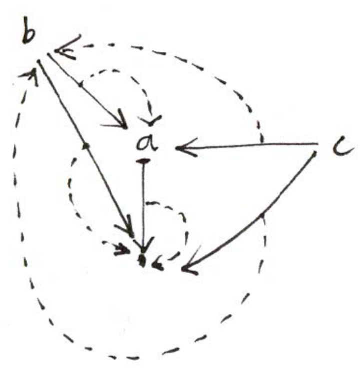
In order to find the power set of $[H,h]$ in the sense of $\mathcal{H}$, we need a set in V containing a representative from each bisimulation class of multigraphs that $\mathcal{H}$ believes to be a subset of $[H,h]$. Before giving a rigorous account we start with an informal discussion of the general approach.

If $[Q,q]$ is a subset of $[H,h]$, we seek to map the hypergraph Q onto a subgraph of H and reconstruct a copy of $[Q,q]$ from the image of this map. This gives a canonical representation for each $[Q,q] \mathrel{\overline{\subset}} [H,h]$, and the collection of these representations is a set since they are all built from a set-sized domain.

Let us explore a naive strategy to build such a map, utilising the recursive property of multiset inclusion. If $[Q,q] \mathrel{\overline{\subset}} [H,h]$, first map q to h. If $[Q_x,x] \mathrel{\overline{\in}} [Q,q]$, then $[Q_x,x]$ is bisimilar to a unique $[H_y,y] \mathrel{\overline{\in}} [H,h]$ so we map x to y. Furthermore

$$\tfrac{[Q,q]}{[Q_x,x]} \mathrel{\overline{\subset}} \tfrac{[H,h]}{[H_y,y]}$$

so we can map $Q(q,x)$ to $H(h,y)$ and iterate this algorithm until we run out of vertices, as $[Q,q]$ is accessible. Unfortunately the map is not quite injective, so we cannot reconstruct $[Q,q]$ from the image. For example consider the following hypergraph G:

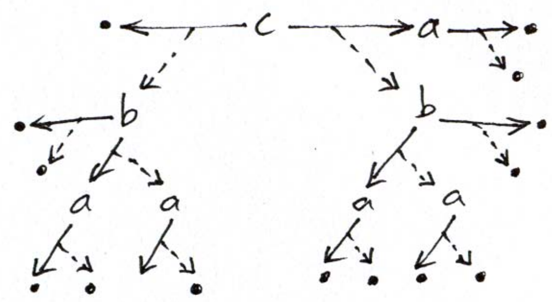


Then

- $\mathcal{H} \models [G_a, a] = \{\emptyset \otimes \emptyset\}$
- $\mathcal{H} \models [G_b, b] = \{\emptyset \otimes \emptyset, [G_a, a] \otimes [G_a, a]\}$
- $\mathcal{H} \models [G_c, c] = \{\emptyset \otimes [G_b, b], [G_a, a] \otimes [G_b, b]\}$

It is clear that $[G_a, a] \mathrel{\overline{\sqsubset}} [G_b, b]$, therefore $[G_b, b] \mathrel{\overline{\sqsubset}} [G_c, c]$. However if we define a map from G_b to G_c following the naive strategy above, then both a and $\bullet$ get mapped to b. The problem is that b is referred to twice in the graph G_c, as the multiplicity of $\emptyset$ and of $[G_a, a]$. Thus to make the map injective we need to make two distinct copies of b and map each of a and the unnamed vertex to a different copy, then repeat the process at lower levels.

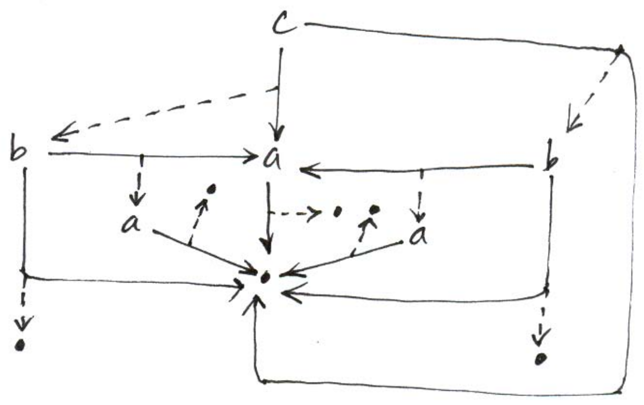
We *unravel* a pointed hypergraph $[H, h]$ by making distinct copies of the same vertex if there are multiple descending path leading to it. For example the hypergraph $[G_c, c]$ above is unraveled as



Note that any copy in our unraveling falls under one of two categories: it is either a member of its parent vertex, or the multiplicity of the parent vertex in the grandparent vertex.

Furthermore, according to our naive strategy earlier the only vertices which can have multiple preimages are multiplicities: roughly speaking since our strategy maps any member of the subset to a bisimilar counterpart in the superset, by extensionality of the subset no two members of the subset can be mapped to the same image.

Thus for any vertex x in the original hypergraph we only need a separate copy for each instance x appears as a multiplicity; all instances of x as a member of other multigraphs can be represented by one copy. This mean we can simplify our earlier unraveling by identifying all member-copies of a vertex. For example the multigraph $[G_c, c]$ can be unraveled as



Note that in the new unraveling each vertex has only one member-copy (the one with solid arrows going in), but possibly more than one multiplicity-copies (with dotted arrows going in). This will help simplify our formal definition of the unraveling.

To keep track of the different multiplicity-copies of the same vertex we index each copy by the path leading down to it. This is always possible since the original hypergraph is accessible, and provides a unique index for each multiplicity-copy. In our example the leftmost copy of the vertex a will be indexed by the path

$$\langle c, a, b \rangle, \langle b, a, a \rangle$$

while the rightmost copy of a will be indexed by

$$\langle c, \bullet, b \rangle, \langle b, a, a \rangle$$

We denote such paths by abbreviated vertex sequences, e.g. (c, a, b, a, a) for the first path and $(c, \bullet, b, a, a)$ for the other. Since only multiplicity-copies are duplicated, adjacent edges in these paths always link up at the 3rd and 1st coordinates; hence no confusion arise from our abbreviation.

Since the member-copy is unique for each vertex, we can just use the original vertex to denote it. In fact it can always be read off from the sequence denoting its corresponding multiplicity-copies, as the next-to-last vertex.

Definition 59. *Write* $\lfloor x_1 \dots x_n \rfloor$ *for the finite sequence* $x_1 \dots x_n$ *where* $n \in \omega$*. Formally we implement* $\lfloor x_1 \dots x_n \rfloor$ *as a function* $n \to V$ *whose value at* i *is* x_{i+1}*.*

To keep our formulae readable, we will sometimes write

$$\forall \lfloor x_1 \dots x_n \rfloor \phi(x_a, x_b, x_c \dots)$$

as shorthand for

$$(\exists n \in \omega)(\exists x : n \to V)\phi(x(a+1), x(b+1), x(c+1) \dots)$$

Definition 60. *The expansion* $\exp[H, h]$ *of* $[H, h] \in \mathcal{H}$ *is the set of all* $\lfloor h_1 \dots h_n \rfloor$ *such that*

- $h_i \in \mathrm{Dom}[H, h]$ *for all* $i \leq n$*.*
- $h_1 = h$*.*
- $h_{i+1} \in H^{-1}h_i$ *for all* i *odd and* $i + 1 \leq n$*.*
- $h_{i+2} = H(h_i, h_{i+1})$ *for all* i *odd and* $i + 2 \leq n$*.*

The expansion of $[H, h]$ formally represents the unraveling of $[H, h]$ by recording all directed paths in said unraveling.

Remark 54. *Given any multigraph* $[H, h] \in \mathcal{H}$*, the axioms of ZF ensure that its expansion exists as a unique set in* V*.*

For example the expansion of the multigraph $[G_c, c]$ in our informal discussion consists of the following sequences, which one can easily match to directed paths in the unraveling of $[G_c, c]$ depicted earlier.

$$\begin{array}{lcr}
\lfloor c \rfloor & \lfloor c, a \rfloor \, \lfloor c, \bullet \rfloor & \lfloor c, a, b \rfloor \\
\lfloor c, \bullet, b \rfloor & \lfloor c, a, b, a \rfloor \, \lfloor c, a, b, \bullet \rfloor & \lfloor c, \bullet, b, a \rfloor \\
\lfloor c, \bullet, b, \bullet \rfloor & \lfloor c, a, b, a, a \rfloor \, \lfloor c, a, b, \bullet, \bullet \rfloor & \lfloor c, \bullet, b, a, a \rfloor \\
\lfloor c, \bullet, b, \bullet, \bullet \rfloor & \lfloor c, a, b, a, a, \bullet \rfloor \, \lfloor c, \bullet, b, a, a, \bullet \rfloor & \lfloor c, a, b, a, a, \bullet, \bullet \rfloor \\
\lfloor c, \bullet, b, a, a, \bullet, \bullet \rfloor & &
\end{array}$$

Remark 55. *Any sequence $\lfloor h_1 \dots h_n \rfloor$ in $\exp[H,h]$ is uniquely determined by the h_i for i even.*

Lemma 79. *$[Q,q] \overline{\sqsubset} [H,h]$ if and only if there is an injection $\phi : \exp[Q,q] \to \exp[H,h]$ such that $\phi\lfloor q \rfloor = \lfloor h \rfloor$ and:*

- *If $\phi\lfloor q_1 \dots q_n \rfloor = \lfloor h_1 \dots h_n \rfloor$ and $\lfloor q_1 \dots q_{n+1} \rfloor \in \exp[Q,q]$, then $\phi\lfloor q_1 \dots q_{n+1} \rfloor = \lfloor h_1 \dots h_{n+1} \rfloor$ for some $h_{n+1} \in \mathrm{Dom}[H,h]$.*
- *If n is even and $\phi\lfloor q_1 \dots q_n \rfloor = \lfloor h_1 \dots h_n \rfloor$, then $[Q_{q_n}, q_n] \cong [H_{h_n}, h_n]$.*

Proof. Let $[Q,q] \overline{\sqsubset} [H,h]$.

Define $\phi\lfloor q_1 \dots q_n \rfloor := \lfloor h_1 \dots h_n \rfloor$ if $[Q_{q_i}, q_i] \overline{\sqsubset} [H_{h_i}, h_i]$ for i odd and $[Qq_i, q_i] \cong [H_{h_i}, h_i]$ for i even.

Clearly $\phi\lfloor a \rfloor = \lfloor h \rfloor$ and by definition if $\phi\lfloor q_1 \dots q_n \rfloor = \lfloor h_1 \dots h_n \rfloor$ where n is even, then $[Q_{q_n}, q_n] \cong [H_{h_n}, h_n]$.

Let $\lfloor q_1 \dots q_n \rfloor \in \exp[Q,q]$. Since $[Q,q] \overline{\sqsubset} [H,h]$, by Lemma 76 there is a finite sequence $h_1 \dots h_n$ in $\mathrm{Dom}[H,h]$ such that:

- $[Q_{q_i}, q_i] \overline{\sqsubset} [H_{h_i}, h_i]$ for i odd.
- $[Qq_i, q_i] \cong [H_{h_i}, h_i]$ for i even.
- $h_1 = h$.
- $h_{i+1} \in H^{-1}h_i$ for i odd and $i+1 \leq n$.
- $h_{i+2} = H(h_i, h_{i+1})$ for i odd and $i+2 \leq n$.

Then $\lfloor h_1 \dots h_n \rfloor \in \exp[H,h]$. Since $[H,h]$ is extensional, h_i is uniquely determined by $[H_{h_i}, h_i]$ for all i even. Therefore $\lfloor h_1 \dots h_n \rfloor$ is unique and ϕ is well-defined.

If $\phi\lfloor q_1 \dots q_n \rfloor = \phi\lfloor b_1 \dots b_m \rfloor$, then $m = n$ by definition of ϕ. Furthermore $[Q_{q_i}, q_i] \cong [Q_{b_i}, b_i]$ for all i even, so $q_i = b_i$ for all i even. But $q_1 = b_1$ trivially and $q_{i+2} = Q(q_i, q_{i+1})$ for any i odd, so $q_i = b_i$ for all i and thus ϕ is injective.

If $\phi\lfloor q_1 \dots q_n \rfloor = \lfloor h_1 \dots h_n \rfloor$ and $\lfloor q_1 \dots q_{n+1} \rfloor \in \exp[Q,q]$, let $\lfloor b_1 \dots b_{n+1} \rfloor = \phi\lfloor q_1 \dots q_{n+1} \rfloor$. By construction $\lfloor b_1 \dots b_n \rfloor = \phi\lfloor q_1 \dots q_n \rfloor$, so $b_i = h_i$ for all $i \leq n$. Setting $h_{n+1} := b_{n+1}$ gives $\phi\lfloor q_1 \dots q_{n+1} \rfloor = \lfloor h_1 \dots h_{n+1} \rfloor$.

Conversely, suppose there exists an injection ϕ satisfying the conditions in the lemma.

Define a relation $\lhd \subset \mathrm{Dom}[Q,q] \times \mathrm{Dom}[H,h]$ as follows:

$x \lhd y$ if there exists $\lfloor q_1 \dots q_n \rfloor \in \exp[Q, q]$ where n is odd and $x = q_n$, such that $\phi\lfloor q_1 \dots q_n \rfloor = \lfloor h_1 \dots h_n \rfloor$ where $y = h_n$.

Since $\phi\lfloor a \rfloor = \lfloor h \rfloor$ we have $a \lhd h$.

Let $\phi\lfloor q_1 \dots q_n \rfloor = \lfloor h_1 \dots h_n \rfloor$ where n is odd. If $q_{n+1} \in Q^{-1}q_n$, let $q_{n+2} = Q(q_n, q_{n+1})$.

Since $\lfloor q_1 \dots q_{n+1} \rfloor, \lfloor q_1 \dots q_{n+2} \rfloor \in \exp[Q, q]$, let $\lfloor h_1 \dots h_{n+1} \rfloor = \phi\lfloor q_1 \dots q_{n+1} \rfloor$ and $\lfloor h_1 \dots h_{n+2} \rfloor = \phi\lfloor q_1 \dots q_{n+2} \rfloor$. Then $[H_{h_{n+1}}, h_{n+1}] \cong [Q_{q_{n+1}}, q_{n+1}]$, $h_{n+2} = H(h_n, h_{n+1})$ and $q_{n+2} \lhd h_{n+2}$.

Hence $\lhd$ witnesses $[Q, q] \mathrel{\overline{\sqsubset}} [H, h]$. □

Now that we have an injection of any $[Q, q] \mathrel{\overline{\sqsubset}} [H, h]$ onto a subgraph of H, all that remains is to reconstruct a copy of $[Q, q]$ from the image.

Lemma 80. *Axiom of Power Set*
If $[H, h] \in \mathcal{H}$, there exists $[D, d] \in \mathcal{H}$ such that

$$(\forall [Q, q] \in \mathcal{H})([Q, q] \mathrel{\overline{\in}} [D, d] \Leftrightarrow [Q, q] \mathrel{\overline{\sqsubset}} [H, h])$$

.

Proof. Say $B \subset \exp[H, h]$ is *good* if

- $\lfloor h \rfloor \in B$.
- If $\lfloor h_1 \dots h_{n+1} \rfloor \in B$, then $\lfloor h_1 \dots h_n \rfloor \in B$.
- If $\lfloor h_1 \dots h_{n+1} \rfloor \in B$ and n is odd, then $\lfloor h_1 \dots h_{n+2} \rfloor \in B$ where $h_{n+2} = H(h_n, h_{n+1})$.

We will show that each good $B \subset \exp[H, h]$ is the image of some multigraph $[Q, q] \mathrel{\overline{\sqsubset}} [H, h]$ under the injection we specified earlier, and each multigraph $[Q, q] \mathrel{\overline{\sqsubset}} [H, h]$ injects onto a unique good $B \subset \exp[H, h]$. First we describe the algorithm to construct a canonical copy of this $[Q, q]$ from B, which we will denote by $[Q^B, q^B]$.

Let ν be an injection where $\operatorname{dom} \nu = \exp[H, h]$ and $\operatorname{ran} \nu$ is disjoint from $\operatorname{Dom}[H, h]$. We will use $\operatorname{ran} \nu$ as the set of vertices on which to construct $[Q^B, q^B]$. Since ν is fixed, this gives a canonical $[Q^B, q^B]$ for all good B.

If $h_1 \dots h_n \in \exp[H, h]$ where n is even, define

$$\begin{aligned} H\lfloor h_1 \dots h_n \rfloor := & H_{h_n} \setminus \{\langle h_n, h_{n+1}, h_{n+2} \rangle : H(h_n, h_{n+1}, h_{n+2})\} \cup \\ & \{\langle \nu\lfloor h_1 \dots h_n \rfloor, h_{n+1}, h_{n+2} \rangle : H(h_n, h_{n+1}, h_{n+2})\} \end{aligned}$$

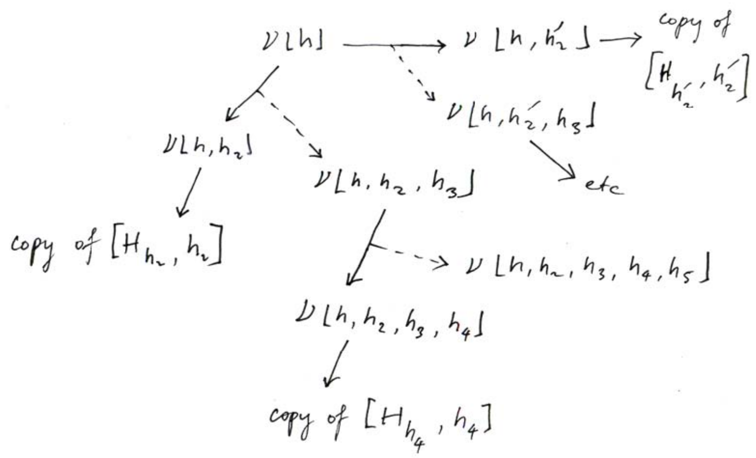
In other words, $[H\lfloor h_1 \dots h_n \rfloor, \nu\lfloor h_1 \dots h_n \rfloor]$ is an isomorphic copy of $[H_{h_n}, h_n]$ obtained by replacing h_n with $\nu\lfloor h_1 \dots h_n \rfloor$.

If B is good, define $\overline{B}$ as the smallest set such that

- If $\lfloor h_1 \dots h_{n+2} \rfloor \in B$ where n is odd, then $\langle \nu\lfloor h_1 \dots h_n \rfloor, \nu\lfloor h_1 \dots h_{n+1} \rfloor, \nu\lfloor h_1 \dots h_{n+2} \rfloor \rangle \in \overline{B}$.

- If $\lfloor h_1 \dots h_n \rfloor \in B$ where n is even, then $H\lfloor h_1 \dots h_n \rfloor \subset \overline{B}$.

To explain the construction, informally $\overline{B}$ is a reconstruction of the unraveled hypergraph represented by B. We built $\overline{B}$ from the top down using the range of ν as the set of vertices. If n is odd, then h_n is the multiplicity of h_{n-1} in h_{n-2} so we made separate copies of h_n in $\overline{B}$ according to the different paths leading to h_n in B. If n is even, then h_n is a member of h_{n-1} so we simply copied $[H_{h_n}, h_n]$ over to $\overline{B}$.



Above is an illustration of the construction. In the illustrated example, since h_3 is the multiplicities of both h_2 and h_2' we made two distinct copies of h_3 in $\overline{B}$, each indexed by the path leading to it from the top vertex h.

By the same reasoning as in the proof of Lemma 77, $[\overline{B}, \nu\lfloor h \rfloor]$ is accessible.

Furthermore if $\lfloor h_1 \dots h_n \rfloor \in B$ and n is even, then by construction

$$[\overline{B}_\nu\lfloor h_1 \dots h_n \rfloor, \nu\lfloor h_1 \dots h_n \rfloor] = [H\lfloor h_1 \dots h_n \rfloor, \nu\lfloor h_1 \dots h_n \rfloor]$$

Suppose $\overline{B}(x, a, v)$ and $\overline{B}(x, b, w)$ where $[\overline{B}_a, a] \equiv [\overline{B}_b, b]$.

If $x = \nu\lfloor h_1 \dots h_n \rfloor$ where n is odd, then $a = \nu\lfloor h_1 \dots h_n, a_{n+1} \rfloor$ and $b = \nu\lfloor h_1 \dots h_n, b_{n+1} \rfloor$ for some $a_{n+1}, b_{n+1} \in \mathrm{Dom}[H, h]$.

But then $[\overline{B}_a, a] \cong [H_{a_{n+1}}, a_{n+1}]$ and $[\overline{B}_b, b] \cong [H_{b_{n+1}}, b_{n+1}]$ by construction, so $[H_{a_{n+1}}, a_{n+1}] \equiv [H_{b_{n+1}}, b_{n+1}]$ and thus $a_{n+1} = b_{n+1}$ since $[H, h]$ is extensional. This means $a = b$, so $v = w$ by construction of $\overline{B}$.

If $x = \nu\lfloor h_1 \dots h_n \rfloor$ where n is odd or $x \in \mathrm{Dom}[H, h]$, then $[\overline{B}_x, x] \in \mathcal{H}$ by construction of $\overline{B}$. But $[\overline{B}_a, a] \equiv [\overline{B}_b, b]$, so $a = b$ and thus $v = w$.

Let $[Q^B, q^B]$ be the extensional quotient of $[\overline{B}, \nu\lfloor h \rfloor]$ with quotient map π, then we have proved that

$$[Q^B, q^B] \in \mathcal{H}$$

We employ Lemma 79 to show that $[Q^B, q^B] \mathrel{\overline{\subset}} [H, h]$:

If $\lfloor q_1 \dots q_n \rfloor \in \exp[Q^B, q^B]$, then $q_{i+1} \in (Q^B)^{-1} q_i$ and $q_{i+2} = Q^B(q_i, q_{i+1})$ for all i odd.

Furthermore $q_1 = q^B$, so by construction of $[Q^B, q^B]$ there exists $\lfloor h_1 \dots h_n \rfloor \in B$ such that $q_i = \pi\nu\lfloor h_1 \dots h_i \rfloor$ for all i.

If i is even, then $[\overline{B}_{\nu\lfloor h_1 \dots h_i \rfloor}, \nu\lfloor h_1 \dots h_i \rfloor] \cong [H_{h_i}, h_i]$ by construction, so $[Q^B_{q_i}, q_i] \cong [H_{h_i}, h_i]$.

Since $[H, h]$ is extensional, h_i is unique for all i even, and thus $\lfloor h_1 \dots h_n \rfloor$ is unique.

Therefore we can define a function $\phi : \exp[Q^B, q^B] \to B$ by setting $\phi\lfloor q_1 \dots q_n \rfloor := \lfloor h_1 \dots h_n \rfloor$.

If $\phi\lfloor q_1 \dots q_n \rfloor = \phi\lfloor a_1 \dots a_n \rfloor$, then $[Q^B_{q_i}, q_i] \cong [Q^B_{a_i}, a_i]$ for all i even. Hence $q_i = a_i$ for all i even, and so $\lfloor q_1 \dots q_n \rfloor = \lfloor a_1 \dots a_n \rfloor$. This shows that ϕ is injective.

If $\lfloor h_1 \dots h_n \rfloor \in B$, let $q_i = \pi\nu\lfloor h_1 \dots h_i \rfloor$ for all i. Then $\lfloor q_1 \dots q_n \rfloor \in \exp[Q^B, q^B]$ by construction of $[Q^B, q^B]$, so ϕ is surjective.

Since ϕ satisfies the conditions in Lemma 79, we have

$$[Q^B, q^B] \mathrel{\overline{\subset}} [H, h]$$

Now given $[Q, q] \mathrel{\overline{\subset}} [H, h]$ we construct a good $B \mathrel{\overline{\subset}} \exp[H, h]$ that encodes $[Q, q]$.

There is an injection $\psi : \exp[Q, q] \to \exp[H, h]$ satisfying the conditions in Lemma 79, so let $B = \operatorname{ran}\psi$. Clearly $\lfloor h \rfloor \in B$.

If $\lfloor h_1 \dots h_{n+1} \rfloor = \psi\lfloor q_1 \dots q_{n+1} \rfloor$, let $\lfloor v_1 \dots v_n \rfloor = \psi\lfloor q_1 \dots q_n \rfloor$. There exists $v_{n+1} \in \mathrm{Dom}[H, h]$ such that $\lfloor v_1 \dots v_{n+1} \rfloor = \psi\lfloor q_1 \dots q_{n+1} \rfloor$, but ψ is injective so $v_i = h_i$ for all i. Hence $\lfloor h_1 \dots h_n \rfloor = \psi\lfloor q_1 \dots q_n \rfloor$.

If $\lfloor h_1 \dots h_{n+1} \rfloor = \psi\lfloor q_1 \dots q_{n+1} \rfloor$ where n is odd, then $\lfloor h_1 \dots h_{n+2} \rfloor = \psi\lfloor q_1 \dots q_{n+2} \rfloor$ where $hn + 2 = H(h_n, h_{n+1})$ and $q_{n+2} = H(q_n, q_{n+1})$.

Therefore B is good, so we can construct $[Q^B, Q^b]$ and $\phi : \exp[Q^B, q^B] \to B$ as earlier in the proof.

It is straightforward to see that the maps $\phi^{-1}\psi : \exp[Q, q] \to \exp[Q^B, q^B]$ and $\psi^{-1}\phi : \exp[Q^B, q^B] \to \exp[Q, q]$ both satisfy the conditions in Lemma 79.

Hence $[Q, q] \mathrel{\overline{\subset}} [Q^B, q^B]$ and $[Q^B, q^B] \mathrel{\overline{\subset}} [Q, q]$, so

$$[Q, q] \cong [Q^B, q^B]$$

This means the set of $[Q^B, q^B]$ for all good $B \mathrel{\overline{\subset}} \exp[H, h]$ contains one representative from each isomorphism class of multigraph $[Q, q] \mathrel{\overline{\subset}} [H, h]$, allowing us to use the supertransitivity of $\mathcal{H}$ to obtain the power set of $[H, h]$ in the sense of $\mathcal{H}$.

Define a constant function δ on $\{[Q^B, q^B] : B \subset \exp[H, h] \wedge B \text{ is good}\}$ by $\delta[Q, q] = [H, h]$.

We have proved that for any $[Q, q] \in \mathcal{H}$

$$[Q, q] \mathrel{\overline{\subset}} [H, h] \Leftrightarrow (\exists [Q^B, q^B] \in \operatorname{dom} \delta)[Q, q] \cong [Q^B, q^B]$$

Therefore by Lemma 78 there exists $[D, d] \in \mathcal{H}$ such that

$$(\forall [Q, q] \in \mathcal{H})([Q, q] \mathrel{\overline{\subset}} [H, h] \Leftrightarrow [Q, q] \mathrel{\overline{\in}} [D, d])$$

. ∎

Lemma 81. *Axiom schema of Multiplicity Replacement*
Let $[H, h] \in \mathcal{H}$ and $\phi(x)$ be a formula in $\mathcal{L}_{\mathcal{H}}$ (see Definition 22) with parameters in $\mathcal{H}$. Suppose for any $[Q, q] \mathrel{\overline{\in}} [H, h]$ there is a unique $[A, a] \in \mathcal{H}$ (up to isomorphism) such that $\mathcal{H} \models \phi([Q, q], [A, a])$. Then there exists $[D, d] \in \mathcal{H}$ such that

$$(\forall [Q, q] \in \mathcal{H})([Q, q] \mathrel{\overline{\in}} [D, d] \Leftrightarrow [Q, q] \mathrel{\overline{\in}} [H, h])$$

and

$$(\forall [Q, q] \mathrel{\overline{\in}} [D, d])\phi([Q, q], \tfrac{[D,d]}{[Q,q]})$$

Proof. By Collection and Comprehension in V, there is a set $M \subset \mathcal{H}$ such that

$$\mathcal{H} \models M = \{[A, a] : (\exists [Q, q] \mathrel{\overline{\in}} [H, h])\phi([Q, q], [A, a])\}$$

Without loss of generality we can arrange for all members of M to have disjoint domains, and that they are all disjoint from $\operatorname{Dom}[H, h]$. We construct a pointed hypergraph $[B, b]$ as follows:

- Let $B := \bigcup\{H_x : x \in H^{-1}h\} \cup \bigcup\{A : [A, a] \in M\}$

- Add to B a new vertex b and edges $\langle b, x, a\rangle$ for each pair $x \in H^{-1}h$ and $[A, a] \in M$ such that $\phi([H_x, x], [A, a])$.

Then $[B, b]$ satisfies the condition in Lemma 72, so its extensional quotient $[D, d]$ is a multigraph. By the Quotient Lemma, it is straightforward to verify that $[D, d]$ is the required multigraph. $\square$

Lemma 82. *Axiom schema of Comprehension*
Let $[H, h] \in \mathcal{H}$ and $\phi(x)$ be a formula in $\mathcal{L}_{\mathcal{H}}$ with parameters in $\mathcal{H}$. Then there exists $[D, d] \in \mathcal{H}$ such that

$$(\forall [Q, q] \in \mathcal{H})([Q, q] \mathrel{\overline{\in}} [D, d] \Leftrightarrow ([Q, q] \mathrel{\overline{\in}} [H, h] \wedge \mathcal{H} \models \phi([Q, q])))$$

and

$$(\forall [Q, q] \mathrel{\overline{\in}} [D, d])(\tfrac{[D,d]}{[Q,q]} \cong \tfrac{[H,h]}{[Q,q]})$$

Proof. Let

$$A := \{[H_a, a] : a \in H^{-1}h \wedge \mathcal{H} \models \phi([H_a, a])\}$$

and define $\delta : A \to \mathcal{H}$ by $\delta([H_a, a]) := \frac{[H,h]}{[H_a,a]}$.

If $[H_a, a], [H_b, b] \in A$ such that $[H_a, a] \cong [H_q, q]$, then $a = q$ by extensionality of $[H, h]$ so $\delta([H_a, a]) = \delta([H_q, q])$.

Hence by Lemma 78 we have $[D, d] \in \mathcal{H}$ as required. $\square$

It is easy to see that with the presence of Multiplicity Replacement, the axiom of Collection below implies the axiom of Replacement as formulated in the theory MS.

Lemma 83. *Axiom of Collection*
Let $[H, h] \in \mathcal{H}$ and $\phi(x, y)$ be a formula in $\mathcal{L}_{\mathcal{H}}$ with parameters in $\mathcal{H}$ such that

$$(\forall [Q, q] \mathrel{\overline{\in}} [H, h])(\exists [A, a] \in \mathcal{H})\mathcal{H} \models \phi([Q, q], [A, a])$$

Then there exists $[D, d] \in \mathcal{H}$ such that

$$(\forall [Q, q] \mathrel{\overline{\in}} [H, h])(\exists [A, a] \mathrel{\overline{\in}} [D, d])\mathcal{H} \models \phi([Q, q], [A, a])$$

Proof. By Collection in V, let B be a set such that

$$(\forall q \in H^{-1}h)(\exists [A, a] \in B)\mathcal{H} \models \phi([H_q, q], [A, a])$$

By Comprehension we have a set

$$\operatorname{dom} \delta := \{[A, a] \in B \cap \mathcal{H} : (\exists [Q, q] \mathrel{\overline{\in}} [H, h])\mathcal{H} \models \phi([Q, q], [A, a])\}$$

Define δ on $\operatorname{dom}\delta$ by $\delta([A,a]) := [H,h]$. By Lemma 78 we have $[D,d] \in \mathcal{H}$ as required. $\square$

Lemma 84. *Axiom of Infinity*
There is a well-ordered multigraph $\boldsymbol{\omega}$ *that satisfies Axiom 17.*

Proof. Let $\alpha_0 := [\emptyset, 0]$, an empty multiset in the sense of $\mathcal{H}$.

If $\alpha_n = [A,a]$, define $\alpha_{n+1} := [B,b]$ where $b \notin \mathrm{Dom}[A,a]$ and

$$B = A \cup \{\langle b, x, 0\rangle : x \in A^{-1}a\} \cup \{b, a, 0\}$$

By induction on n, it is straightforward to show that $\alpha_n \mathrel{\overline{\in}} \mathcal{H}$ and

$$\mathcal{H} \models \alpha_{n+1} = \alpha_n \cup \{\alpha_n \otimes \boldsymbol{\emptyset}\}$$

Define a function $\phi : \{\alpha_n : n \in \omega\} \to \mathcal{H}$ by $\phi(\alpha_n) = [\emptyset, 0]$, then Lemma 78 gives us the required multigraph. $\square$

We have shown that the given interpretation of the language of multisets in $\mathcal{H}$ produces a model of MS^+, whence

Theorem 7. *If ZF is consistent, then* MS^+ *is consistent.*

Furthermore within the theory MS^+ we can implement accessible pointed hypergraphs and define bisimilarity just like in ZF. Then our model satisties the following anti-foundation axiom in the language of multisets.

Axiom 24. *Multiset AFA*
If H *is a hypergraph such that*

$$(\forall a,b,c,d,e \mathrel{\overline{\in}} \mathrm{Dom}\, H)\Big((H(a,b,c) \wedge H(a,d,e) \wedge [H_b, b] \equiv [H_d, d]) \Rightarrow [H_c, c] \equiv [H_e, e]\Big)$$

then there exists a unique function ϕ *such that* $\operatorname{dom}\phi = \mathrm{Dom}\, H$ *and*

$$(\forall x \mathrel{\overline{\in}} \mathrm{Dom}\, H)(\forall a, b)\Big(a \in_b x \Leftrightarrow (\exists \langle x,y,z\rangle \mathrel{\overline{\in}} H)(\phi(y) = a \wedge \phi(z) = b)\Big)$$

Lemma 85. *Under the given interpretation,* $\mathcal{H}$ *is a model of multiset AFA.*

Proof. Suppose $\mathcal{H}$ believes $[H,h]$ is a hypergraph satisfying the condition of multiset AFA. Define a hypergraph G as follows

$$(\forall x,y,z \in \mathrm{Dom}[H,h])(\langle x,y,z\rangle \in G \Leftrightarrow \mathcal{H} \models \langle [H_x, x], [H_y, y], [H_z, z]\rangle \mathrel{\overline{\in}} [H,h])$$

It is straightforward to verify that $[G_x, x]$ satisfies the conditions of Lemma 72 for any $x \in \operatorname{Dom} G$, so its extensional quotient is a multigraph.

Moreover for all $x, y, z \in \operatorname{Dom} G$

$$(\mathcal{H} \models [G_y, y] \in_{[G_z,z]} [G_x, x]) \Leftrightarrow (\mathcal{H} \models \langle [H_x, x], [H_y, y], [H_z, z] \rangle \mathrel{\overline{\in}} [H, h])$$

For each $x \in \operatorname{Dom} G$ we can define a canonical $M_x \in \mathcal{H}$ such that

$$\mathcal{H} \models M_x = \langle [H_x, x], [G_x, x] \rangle$$

Define a constant function $\psi : \{M_x : x \in \operatorname{Dom} G\} \to \mathcal{H}$ by $\phi(M_x) = [\emptyset, 0]$. Then the supertransitivity lemma (Lemma 78) gives the required object in $\mathcal{H}$. $\square$

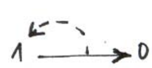
3.2.1 A model where the inclusion relation is not antisymmetric

In any well-founded model of MS an induction on the recursive definition of $\overline{\subset}$ will show it to be antisymmetric, hence the axiom of Foundation implies the axiom of Subset. However as an application of the anti-foundation property of our model $\mathcal{H}$, we will modify it slightly to obtain a model of MS$^+$ where $\overline{\subset}$ is not antisymmetric. Thus the axiom of Subset is not redundant if we want to enforce antisymmetry of the inclusion relation, since it does not follow from the other axioms of MS$^+$.

Consider the semi-Quine atom x where $\emptyset \in_x x$ and x has no other member in the sense of $\mathcal{H}$. If there are two distinct semi-Quine atoms x and y, then Extensionality still holds since the multiplicity of $\emptyset$ is different in x and y. On the other hand it is easy to check that $x \mathrel{\overline{\subset}} y$ and $y \mathrel{\overline{\subset}} x$, so $\overline{\subset}$ is no longer antisymmetric. Although the model $\mathcal{H}$ contains only one semi-Quine atom, we will create one more by dividing its equivalence class into two.

Let $\mathcal{H}$ be the class of hypergraphs as defined in Definition 53. In this section we will redefine multigraphs by strengthening the notions of a bisimulation as follows:

Definition 61. *Let $\mathbb{A} := [\langle 1, 0, 1 \rangle, 1]$, i.e. the pointed hypergraph depicted below.*



Definition 62. *A relation $\sim \subset \operatorname{Dom}[G, g] \times \operatorname{Dom}[H, h]$ is a bisimulation between $[G, g]$ and $[H, h]$ if all of the following hold:*

- *It is a bisimulation in the old definition, i.e. Definition 49.*

- *If* $1 \sim x$ *and* $[G_1, 1] = \mathbb{A}$*, then* $1 \in \mathrm{Dom}[H_x, x]$ *and* $[H_1, 1] = \mathbb{A}$*.*
- *If* $x \sim 1$ *and* $[H_1, 1] = \mathbb{A}$*, then* $1 \in \mathrm{Dom}[G_x, x]$ *and* $[G_1, 1] = \mathbb{A}$*.*

Informally speaking, under this new definition $\mathbb{A}$ is no longer equivalent to its other isomorphic images. Therefore the former equivalence class of any hypergraph containing a semi-Quine atom is now divided into many smaller equivalence classes.

Using the result of Lemma 68 for the old definition of bisimulation, it is trivial to show that the same lemma holds for the new definition since we only need to check the extra clauses. We restate the Lemma below for ease of reference.

Lemma 86.

i If $\sim$ *is a bisimulation between* $[G, g]$ *and* $[H, h]$*, then the relation* $x \simeq y \Leftrightarrow_{df} y \sim x$ *is a bisimulation between* $[H, h]$ *and* $[G, g]$*.*

ii If $\sim$ *is a bisimulation between* $[G, g]$ *and* $[H, h]$ *and* $\simeq$ *a bisimulation between* $[H, h]$ *and* $[Q, q]$*, then the relation* $x \approx d \Leftrightarrow_{df} (\exists a)(x \sim a \wedge a \simeq d)$ *is a bisimulation between* $[G, g]$ *and* $[Q, q]$*.*

iii If $\sim$ *is a bisimulation between* $[G, g]$ *and* $[H, h]$*, then its restriction to* $\mathrm{Dom}[G_x, x] \times \mathrm{Dom}[H_y, y]$ *is a bisimulation between* $[G_x, x]$ *and* $[H_y, y]$*.*

iv Any bisimulation between $[G_x, x]$ *and* $[H_y, y]$ *is a bisimulation between* $[G, g]$ *and* $[H, h]$*.*

v Let $\sim$ *be a bisimulation between* $[G, g]$ *and* $[H, h]$ *such that* $g \sim h$*. If* $[G, g]$ *and* $[G, h]$ *are accessible, then for any* $y \in \mathrm{Dom}[H, h]$ *there is* $x \in \mathrm{Dom}[G, g]$ *such that* $x \sim y$*, and for any* $x \in \mathrm{Dom}[G, g]$ *there is* $y \in \mathrm{Dom}[H, h]$ *such that* $x \sim y$*.*

Definition 63. *Say* $[H, h]$ *is extensional if any bisimulation on* $[H, h]$ *is the identity.*

The proof of Lemma 69 depends only on the truth of Lemma 68, so it also carries over to the new definition trivially and we have the following result.

Lemma 87.

i If $[H, h]$ *is extensional, then so is* $[H_x, x]$ *for any* $x \in \mathrm{Dom}[H, h]$*.*

ii If $[H, h]$ *is extensional and* $[H_x, x] \cong [H_y, y]$ *for* $x, y \in \mathrm{Dom}[H, h]$*, then* $x = y$*.*

From this point on we construct the model in exactly the same way as before except for a minor subtlety. Firstly we redefine the class $\mathcal{H}$ of multigraphs using the new definition of bisimulation, and define relations on $\mathcal{H}$ to stand in for the identity and membership relations:

Definition 64. *Say $[G, g] \equiv [H, h]$, i.e. they are bisimilar, if there is a bisimulation $\sim$ between them such that $g \sim h$.*

Definition 65. *Say $[G, g] \mathrel{\overline{\in}} [H, h]$ if there exists $x \in \mathrm{Dom}[H, h]$ such that $[G, g] \equiv [H_x, x]$.*

If $[G, g] \equiv [H, h]$, then it is easy to see that their extensional quotients are isomorphic. However unlike in the old model, even isomorphic multigraphs need not be bisimilar . For example let $[G, g]$ be a distinct but isomorphic copy of $\mathbb{A}$, then their extensional quotients are just themselves but there is no bisimulation between them that relates the top vertices to each other (since by Definition 62 we would then have $g = 1$ and $[G, g] = \mathbb{A}$). Nevertheless we can easily check that $\equiv$ is an equivalence relation that respects $\overline{\in}$ using Lemma 86. Thus when reproving previous results in the new model we need to replace all instances of isomorphism with the bisimilarity relation.

Remark 56. *We could have developed the previous model in the same way, using the bisimilarity relation instead of the isomorphism relation, since the two relations are the same for extensional pointed hypergraphs. If we had followed that approach, much of the cosmetic changes in the following section could have been avoided. It is however my personal opinion that using the isomorphism relation to interpret the identity relation for multisets is more intuitively clear, thus I have chosen to use the isomorphism relation in the previous section for the sake of a better exposition. In any case, most of the changes that arise from that choice are conceptually trivial.*

The Quotient Lemma requires a bit of care. If we simply use the extensional quotient given under the old definition, we may collapse more than allowed under the new definition. Instead we mirror the proof of the old Quotient Lemma.

Lemma 88. *For any pointed hypergraph $[H, h]$, there exists an extensional pointed hypergraph $[Q, q]$ and a surjective quotient map $\pi : \mathrm{Dom}[H, h] \to \mathrm{Dom}[Q, q]$ such that $q = \pi(h)$, $(\forall a, b, c \in \mathrm{Dom}[Q, q](Q(a, b, c) \Leftrightarrow (\exists x, y, z \in \mathrm{Dom}[H, h])(a = \pi(x) \wedge b = \pi(y) \wedge c = \pi(z)))$, and the relation $\pi(x) = y$ is a bisimulation between $[H, h]$ and $[Q, q]$.*

Furthermore:

i $[Q, q]$ is the unique hypergraph satisfying the conditions above.

ii $\mathrm{Dom}[Q_\pi(x), \pi(x)] = \{\pi(y) : y \in \mathrm{Dom}[H_x, x]\}$ *for any* $x \in \mathrm{Dom}[H, h]$*. In particular if* $[H, h]$ *is accessible, then so is* $[Q, q]$*.*

iii For any $x \in \mathrm{Dom}[H, h]$*, the extensional quotient of* $[H_x, x]$ *given by the first part of the lemma is precisely* $[Q_\pi(x), \pi(x)]$*.*

iv If $[H_x, x]$ *is extensional, then* $[Q_\pi(x), \pi(x)] \cong [H_x, x]$*.*

Proof. First we define the extensional quotient and show that the quotient map is a bisimulation. If $1 \notin \mathrm{Dom}[H, h]$ or $[H_1, 1] \neq \mathbb{A}$, then the new definition of bisimulation on $[H, h]$ reduces to the old definition and we proceed exactly as in Lemma 70. Therefore we only need to consider the case where $[H_1, 1] = \mathbb{A}$.

For $\sim \subset \mathrm{Dom}[H, h]^2$ we redefine $\sim^+$ as follows

$$\begin{aligned} a \sim^+ x \Leftrightarrow_{df} & (\forall b, c \in \mathrm{Dom}[H, h])(H(a, b, c) \Rightarrow (\exists y, z \in \mathrm{Dom}[H, h])(H(x, y, z) \wedge b \sim y \wedge c \sim z)) \wedge \\ & (\forall y, z \in \mathrm{Dom}[H, h])(H(x, y, z) \Rightarrow (\exists b, c \in \mathrm{Dom}[H, h])(H(a, b, c) \wedge b \sim y \wedge c \sim z)) \wedge \\ & [H_a, a] = \mathbb{A} \Rightarrow a \in \mathrm{Dom}[H_b, b] \wedge \\ & [H_b, b] = \mathbb{A} \Rightarrow b \in \mathrm{Dom}[H_a, a] \end{aligned}$$

It is simple to check that $\sim_1 \subset \sim_2$ implies $\sim_1^+ \subset \sim_2^+$ and $\sim$ is a bisimulation if and only if $\sim \subset \sim^+$. Since the redefined operation is still monotonic, as before by the Knaster-Tarski theorem the following relation

$$x \approx y \Leftrightarrow_{df} (\exists \sim)(\sim \subset \sim^+ \wedge x \sim y)$$

is the greatest fixed point, i.e. $\approx = \approx^+$.

Using Lemma 86, it is easy to show that $\approx$ is an equivalence relation, and note that the equivalence classes of 0 and 1 are singletons.

As usual we let $\mathrm{Dom}[Q, q]$ be the set of equivalence classes of $\approx$, q the equivalence class of h and let $\pi : \mathrm{Dom}[H, h] \to \mathrm{Dom}[Q, q]$ the corresponding quotient map.

However, due to the restricted definition of bisimulation we now replace the equivalence classes of 0 and 1 by 0 and 1 respectively, i.e. let $\pi(0) = 0$ and $\pi(1) = 1$. Note that π is still bijective since previously $0 \notin \operatorname{ran} \pi$ and $1 = \{0\} = \pi(0)$.

Define the relation Q on $\mathrm{Dom}[Q, q]$ by

$$Q(a, b, c) \Leftrightarrow_{df} (\exists x, y, z \in \mathrm{Dom}[H, h])(H(x, y, z) \wedge \pi(x) = a \wedge \pi(y) = b \wedge \pi(z) = c)$$

As in the proof of Lemma 70, if $\pi(x) = a$ and $Q(a, b, c)$, then there are $y, z \in \mathrm{Dom}[H, h]$ such that $H(x, y, z)$ and $\pi(y) = b$, $\pi(z) = c$. Conversely if $H(x, y, z)$, then $Q(\pi(x), \pi(y), \pi(z))$.

Since π fixes 0 and 1 by construction, we have shown the relation $\pi(x) = y$ is a bisimulation with the updated definition.

If $\sim_Q$ is a bisimulation on $[Q, q]$, define a relation $\sim_H$ on $\mathrm{Dom}[H, h]$ by

$$x \sim_H y \Leftrightarrow_{df} \pi(x) \sim_Q \pi(y)$$

Since $\pi(x) = y$ is a bisimulation, Lemma 86 shows that $\sim_H$ is a bisimulation on $[H, h]$. This means $\sim_H \subset \approx$ so

$$(\forall x, y \in \mathrm{Dom}[H, h])(x \sim_H y \Rightarrow \pi(x) = \pi(y))$$

and thus $\sim_Q$ is the identity. We have shown that $[Q, q]$ is extensional.

The rest of the proof proceed exactly as in Lemma 70, using Lemma 86 in place of Lemma 68. $\square$

We sketch a proof of the basic extensionality result for our new model, from which the axiom of Extensionality trivially follows.

Lemma 89. *Let* $[G, g], [H, h] \in \mathcal{H}$. *Suppose*

$$(\forall \langle g, x, y\rangle \in G)(\exists \langle h, a, b\rangle \in H)([G_x, x] \equiv [H_a, a] \wedge [G_y, y] \equiv [H_b, b])$$

and

$$(\forall \langle h, a, b\rangle \in H)(\exists \langle g, x, y\rangle \in G)([G_x, x] \equiv [H_a, a] \wedge [G_y, y] \equiv [H_b, b])$$

Then $[G, g] \equiv [H, h]$.

Proof. Define a relation $\sim \subset \mathrm{Dom}[G, g] \times \mathrm{Dom}[H, h]$ by

$$x \sim y \Leftrightarrow_{df} (x = g \wedge y = h) \vee [G_x, x] \equiv [H_y, y]$$

We prove that $\sim$ is a bisimulation between $[G, g]$ and $[H, h]$.

Suppose $x \sim a$ and $G(x, y, z)$.

If $x = g$ and $a = h$, by the hypothesis there are $b, c \in \mathrm{Dom}[H, h]$ such that $H(a, b, c)$ and $[H_b, b] \equiv [G_y, y]$, $[H_c, c] \equiv [G_z, z]$. Then $y \sim b$ and $z \sim c$ as required.

If $[G_x, x] \equiv [H_a, a]$, there is a bisimulation that relates x to a. By definition of bisimulation there are $b, c \in \mathrm{Dom}[H, h]$ such that $y \equiv b$, $z \equiv c$ and $H(a, b, c)$ as required.

If $[G_x, x] = \mathbb{A}$ and $x \neq g$, then $x \in \mathrm{Dom}[G_v, v]$ for some $v \in G^{-1}g$ or $G^{-2}g$. In either case there exists $p \in \mathrm{Dom}[H, h]$ such that $[G_v, v] \equiv [H_p, p]$, so by definition of bisimulation $1 \in \mathrm{Dom}[H, h]$ and $[H_1, 1] = \mathbb{A}$.

If $[G, g] = \mathbb{A}$, then $G(g, 0, 1)$ and we proceed exactly as in the previous case.

The other direction is similar, so $\sim$ is a bisimulation. $\square$

Let us now address a small inconvenience. In the old model, given any collection of multigraphs we can easily create a copy in which all the multigraphs have disjoint domains. This was crucial in the proof of many axioms such as Collection, where a large multigraph needs to be created. However the special status of $\mathbb{A}$ in our new model means that we are longer allowed to replace it by isomorphic copies. Thus we have to slightly weaken the conditions of disjoint domains.

Definition 66. *If $A \subset \mathcal{H}$ is a collection of multigraphs, say A is almost disjoint if the following hold for any $[G, g], [H, h] \in A$:*

- *if $1 \in \mathrm{Dom}[G, g], 1 \in \mathrm{Dom}[H, h]$ and $[G_1, 1] = [H_1, 1] = \mathbb{A}$, then $\mathrm{Dom}[G, g] \cap \mathrm{Dom}[H, h] = \{0, 1\}$;*
- *otherwise $\mathrm{Dom}[G, g] \cap \mathrm{Dom}[H, h] = \emptyset$.*

Say B is an almost disjoint copy of A if B is almost disjoint and there is a bijection $\pi : A \leftrightarrow B$ such that $[G, g] \equiv \pi[G, g]$ for all $[G, g] \in A$.

Note that this definition also works for classes if we allow the bijection to be a function-class, but for our current purposes it suffices to consider only sets.

Lemma 90. *For any collection $A \subset \mathcal{H}$ there is an almost disjoint copy.*

Proof. If $1 \notin \mathrm{Dom}[G, g]$ or $[G_1, 1] \neq \mathbb{A}$, define $\pi[G, g]$ as follows:

For any $a \in \mathrm{Dom}[G, g]$ let $\hat{a} := \langle a, [G, g] \rangle$, then trivially $a \neq 0$ and $a \neq 1$.

Let $\hat{G} := \{\langle \hat{a}, \hat{b}, \hat{c} \rangle : \langle a, b, c \rangle \in G\}$ and $\pi[G, g] := [\hat{G}, \hat{g}]$. Then $\pi[G, g] \cong [G, g]$ and the isomorphism is a bisimulation between them.

If $1 \in \mathrm{Dom}[G, g]$ and $[G_1, 1] = \mathbb{A}$, define $\hat{a}$ for $a \in \mathrm{Dom}[G, g]$ as above except that $\hat{0} = 0$ and $\hat{1} = 1$ and define $\pi[G, g]$ as above.

Then $\pi[G, g] \cong [G, g]$ and the isomorphism is a bisimulation since both 0 and 1 are fixed.

It is straightforward to check that the range of π is almost disjoint. $\square$

Remark 57. *Let A be an almost disjoint set of multigraphs, and let H be the hypergraph obtained by taking the union of all hypergraphs in A. Then for any $[G, g] \in A$, $[G, g] = [H_g, g]$.*

The counterpart to Lemma 72 can be proved with exactly the same proof as before, so we only state the result below.

Lemma 91. *Let $[H, h]$ be accessible and*

$$(H(a, b, c) \wedge H(a, d, e) \wedge [H_b, b] \equiv [H_d, d]) \Rightarrow [H_c, c] \equiv [H_e, e]$$

for all $a, b, c, d, e \in \mathrm{Dom}[H, h]$. Then the extensional quotient $[Q, q]$ of $[H, h]$ is a multigraph.

Now the new supertransitivity lemma can be proved in exactly the same way as Lemma 78.

Lemma 92. *Let $\phi : \mathcal{H} \to \mathcal{H}$ be a function in the ZF model V such that*

$$(\forall [G, g], [H, h] \in \mathrm{dom}\, \phi)[G, g] \equiv [H, h] \Rightarrow \phi[G, g] = \phi[H, h]$$

Then there exists $[Q, q] \in \mathcal{H}$ such that

$$(\forall [H, h] \in \mathcal{H})\Big([H, h] \mathrel{\overline{\in}} [Q, q] \Leftrightarrow (\exists [G, g] \in \mathrm{dom}\, \phi)[G, g] \equiv [H, h]\Big)$$

and

$$(\forall [H, h] \in \mathrm{dom}\, \phi)\tfrac{[Q,q]}{[H,h]} \equiv \phi[H, h]$$

Proof. By taking a copy if necessary, assume without loss of generality that $\mathrm{dom}\, \phi \cup \mathrm{ran}\, \phi$ is almost disjoint. The rest of the proof follows exactly as with Lemma 78, replacing the isomorphism relation $\cong$ with $\equiv$. $\square$

Having proved these preliminary results, we can carry over the proofs of Lemmata 82, 81, 83 and 84 to obtain the axioms of Comprehension, Multiplicity Replacement, Collection and Infinity.

Next we turn to the inclusion relation and related axioms. First we give the updated definition of $\overline{\subset}$ and outline a proof of its basic properties.

Definition 67. *Say $[G, g] \mathrel{\overline{\subset}} [H, h]$ if there is a relation $\lhd \subset \mathrm{Dom}[G, g] \times \mathrm{Dom}[H, h]$ such that $g \lhd h$ and*

$$x \lhd y \Rightarrow (\forall a \in G^{-1}x)(\exists b \in H^{-1}y)([G_a, a] \equiv [H_b, b] \wedge G(x, a) \lhd H(y, b))$$

for all $x \in \mathrm{Dom}[G, g]$, $y \in \mathrm{Dom}[H, h]$.

Lemma 93. *If $[G,g] \equiv [H,h]$, $[P,p] \equiv [Q,q]$ and $[G,g] \mathrel{\overline{\subset}} [P,p]$, then $[H,h] \mathrel{\overline{\subset}} [Q,q]$. Furthermore $\overline{\subset}$ is reflexive, transitive and*

$$[G,g] \mathrel{\overline{\subset}} [H,h] \Leftrightarrow (\forall [Q,q] \mathrel{\overline{\in}} [G,g]) \left([Q,q] \mathrel{\overline{\in}} [H,h] \wedge \tfrac{[G,g]}{[Q,q]} \mathrel{\overline{\subset}} \tfrac{[H,h]}{[Q,q]} \right)$$

Proof. For the first part let $\sim_1, \sim_2$ be the bisimulations involved, and let $\lhd$ witness $[G,g] \mathrel{\overline{\subset}} [P,p]$. Then the relation $\prec$ defined by

$$x \prec y \Leftrightarrow_{df} (\exists a \in \mathrm{Dom}[G,g])(\exists b \in \mathrm{Dom}[P,p])(a \sim_1 x \wedge b \sim_2 y \wedge a \lhd b)$$

witnesses that $[H,h] \mathrel{\overline{\subset}} [Q,q]$ by a straightforward verification against the definition.

The rest of the proof is a straightforward adaptation of the arguments used for Lemma 76, replacing isomorphism with bisimilarity where applicable. □

Note that we can no longer prove the antisymmetry of the inclusion relation. In Lemma 76 we constructed a bisimulation explicitly to show that $\overline{\subset}$ is antisymmetric, but that construction fails to be a bisimulation in the new definition. In fact the purpose of this new model is to break the antisymmetry of the inclusion relation:

Lemma 94. $(\exists [G,g],[H,h] \in \mathcal{H})([G,g] \mathrel{\overline{\subset}} [H,h] \wedge [H,h] \mathrel{\overline{\subset}} [G,g] \wedge [G,g] \not\equiv [H,h])$

Proof. Let $[G,g] = \mathbb{A}$ and $[H,h]$ be a distinct but isomorphic copy. Since the empty multigraph $[\emptyset, 0]$ is equivalent to any isomorphic copy of itself, the isomorphism witnesses both $[G,g] \mathrel{\overline{\subset}} [H,h]$ and $[H,h] \mathrel{\overline{\subset}} [G,g]$ but clearly $[G,g] \not\equiv [H,h]$ (since by Definition 62 we would have $h = 1$ and $[H,h] = \mathbb{A}$). □

Next is the axiom of Union.

Lemma 95. *For any $[H,h] \in \mathcal{H}$ the collection of $[G,g] \in \mathcal{H}$ such that $[G,g] \mathrel{\overline{\in}} [H,h]$ has a least upper bound with respect to $\overline{\subset}$.*

Proof. We mirror the proof of Lemma 77, replacing $\cong$ with $\equiv$ throughout.

The only part of the proof that does not carry over is the uniqueness of a if $a \prec X$ for a given X. As with the antisymmetry of $\overline{\subset}$, that part of the proof explicitly constructs a bisimulation which does not qualify under the new definition. However this uniqueness is not necessary for the rest of the argument. □

We reuse Definition 60 and state the following result, which has the same proof as Lemma 79.

Lemma 96. $[Q,q] \mathrel{\overline{\subset}} [H,h]$ *if and only if there is an injection* $\phi : \exp[Q,q] \to \exp[H,h]$ *such that* $\phi\lfloor q \rfloor = \lfloor h \rfloor$ *and:*

- *If* $\phi\lfloor q_1 \dots q_n \rfloor = \lfloor h_1 \dots h_n \rfloor$ *and* $\lfloor q_1 \dots q_{n+1} \rfloor \in \exp[Q,q]$*, then* $\phi\lfloor q_1 \dots q_{n+1} \rfloor = \lfloor h_1 \dots h_{n+1} \rfloor$ *for some* $h_{n+1} \in \mathrm{Dom}[H,h]$*.*
- *If* n *is even and* $\phi\lfloor q_1 \dots q_n \rfloor = \lfloor h_1 \dots h_n \rfloor$*, then* $[Q_{q_n}, q_n] \equiv [H_{h_n}, h_n]$*.*

We are now in a position to prove the axiom of Power Set.

Lemma 97. *If* $[H,h] \in \mathcal{H}$*, there exists* $[D,d] \in \mathcal{H}$ *such that for any* $[Q,q] \in \mathcal{H}$*,* $[Q,q] \mathrel{\overline{\in}} [D,d]$ *if and only if* $[Q,q] \mathrel{\overline{\subset}} [H,h]$*.*

Proof. We adapt the construction in the proof of Lemma 80 by making provisions for the updated definition of bisimulation.

By replacing $[H,h]$ with a bisimilar copy if necessary, we can assume without loss of generality that either $\mathrm{Dom}[H,h] \cap \{0,1\} = \emptyset$ or $[H_1, 1] = \mathbb{A}$.

Let μ be a bijection such that $\operatorname{dom}\mu = \exp[H,h]$ but $\operatorname{ran}\mu$ is disjoint from both $\mathrm{Dom}[H,h]$ and $\{0,1\}$.

Suppose $[Q,q] \mathrel{\overline{\subset}} [H,h]$. Then there is an injection $\phi : \exp[Q,q] \to \exp[H,h]$ satisfying the conditions in the lemma above.

Define a function ν on $\operatorname{ran}\phi$ by $\nu\phi\lfloor q_1 \dots q_n \rfloor = \mu\phi\lfloor q_1 \dots q_n \rfloor$ if $[Q_{q_n}, q_n] \neq \mathbb{A}$ and $\nu\phi\lfloor q_1 \dots q_n \rfloor = 1$ if $[Q_{q_n}, q_n] = \mathbb{A}$.

If $\lfloor h_1 \dots h_n \rfloor \in \exp[H,h]$ where n is even, define

$$\begin{aligned} H^{\lfloor h_1 \dots h_n \rfloor} :=& H_{h_n} \setminus \{\langle h_n, h_{n+1}, h_{n+2}\rangle : H(h_n, h_{n+1}, h_{n+2})\} \cup \\ & \{\langle \nu\lfloor h_1 \dots h_n \rfloor, h_{n+1}, h_{n+2}\rangle : H(h_n, h_{n+1}, h_{n+2})\} \end{aligned}$$

In other words, $[H^{\lfloor h_1 \dots h_n \rfloor}, \nu\lfloor h_1 \dots h_n \rfloor]$ is an isomorphic copy of $[H_{h_n}, h_n]$ obtained by replacing h_n with $\nu\lfloor h_1 \dots h_n \rfloor$.

One can easily check by the properties of ϕ as stated in the previous lemma that for n even

$$[H^{\phi\lfloor q_1 \dots q_n \rfloor}, \nu\phi\lfloor q_1 \dots q_n \rfloor] \equiv [Q_{q_n}, q_n]$$

Define a hypergraph D as the smallest set such that:

- If $\lfloor q_1 \dots q_{n+2} \rfloor \in \exp[Q,q]$ and n is odd, then

$$\langle \nu\phi\lfloor q_1 \dots q_n \rfloor, \nu\phi\lfloor q_1 \dots q_{n+1} \rfloor, \nu\phi\lfloor q_1 \dots q_{n+2} \rfloor\rangle \in D$$

- If $\lfloor q_1 \dots q_n \rfloor \in \exp[Q, q]$ where n is even, then $H^{\phi \lfloor q_1 \dots q_n \rfloor} \subset D$.
- If $1 \in \operatorname{ran} \nu$, then $\langle 1, 0, 1 \rangle \in D$.

Let $d := \nu\phi\lfloor q \rfloor$, then $[D, d]$ is accessible by construction of D.

Define a relation $\sim \subset \mathrm{Dom}[Q, q] \times \mathrm{Dom}[D, d]$ by

$$x \sim y \Leftrightarrow_{df} [Q_x, x] \equiv [D_y, y] \vee (\exists \lfloor q_1 \dots q_n \rfloor \in \exp[Q, q])(x = q_n \wedge y = \nu\phi\lfloor q_1 \dots q_n \rfloor$$

We show that $\sim$ is a bisimulation.

Suppose $x \sim y$.

If $[Q_x, x] \equiv [D_y, y]$, by definition of bisimilarity

$$(\forall a, b)(Q(x, a, b) \Rightarrow (\exists v, w)(D(y, v, w) \wedge [D_v, v] \equiv [Q_a, a] \wedge [D_w, w] \equiv [Q_b, b]))$$

and vice versa.

If $y = \nu\phi\lfloor q_1 \dots q_n \rfloor$ and $x = q_n$, by construction of D we have:

- if n is odd, then

  $$(\forall a, b)(Q(x, a, b) \Rightarrow D(y, \nu\phi\lfloor q_1 \dots q_n, a \rfloor, \nu\phi\lfloor q_1 \dots q_n, a, b \rfloor)$$

  and

  $$(\forall v, w)(D(y, v, w) \Rightarrow (\exists a, b)(Q(x, a, b) \wedge v = \nu\phi\lfloor q_1 \dots q_n, a \rfloor \wedge w = \nu\phi\lfloor q_1 \dots q_n, a, b \rfloor))$$

- if n is even, then

  $$[D_y, y] = [H^{\phi \lfloor q_1 \dots q_n \rfloor}, \nu\phi\lfloor q_1 \dots q_n \rfloor] \equiv [Q_{q_n}, q_n] = [Q_x, x]$$

  so we are back to the case where $[D_y, y] \equiv [Q_x, x]$.

We have proved that

$$(\forall a, b \in \mathrm{Dom}[Q, q])(Q(x, a, b) \Rightarrow (\exists v, w \in \mathrm{Dom}[D, d])(D(y, v, w) \wedge a \sim v \wedge b \sim w)) \wedge$$
$$(\forall v, w \in \mathrm{Dom}[D, d])(D(y, v, w) \Rightarrow (\exists a, b \in \mathrm{Dom}[Q, q])(Q(x, a, b) \wedge a \sim v \wedge b \sim w))$$

Now we show that

$$[Q_x, x] = \mathbb{A} \Rightarrow (1 \in \mathrm{Dom}[D_y, y] \wedge [D_1, 1] = \mathbb{A})$$

and

$$[D_y, y] = \mathbb{A} \Rightarrow (1 \in \mathrm{Dom}[Q_x, x] \wedge [Q_1, 1] = \mathbb{A})$$

If $[Q_x, x] \equiv [D_y, y]$ the claim is trivial, so we need only consider the case where $y = \nu\phi\lfloor q_1 \dots q_n \rfloor$ and $x = q_n$ for n odd.

If $[Q_x, x] = \mathbb{A}$, then $y = 1$ by definition of ν and so $[D_y, y] = \mathbb{A}$ by the last clause in the construction of D.

If $[D_y, y] = \mathbb{A}$, then $\nu\phi\lfloor q_1 \dots q_n \rfloor = 1$ so by definition of ν it must be that $[Q_x, x] = \mathbb{A}$.

Therefore $\sim$ is a bisimulation. It is easy to show that the extensional quotient of $[D, d]$ is a multigraph, which is then bisimilar to $[Q, q]$.

Let M be the set of all accessible pointed hypergraphs with vertices in ran $\mu \cup \mathrm{Dom}[H, h]$, and let N be the set of extensional quotients of members of M. We have shown that every $[Q, q] \mathrel{\overline{\subset}} [H, h]$ must be bisimilar to a member of N.

Thus we can use the updated supertransitivity lemma to obtain the required multigraph from the set $\{[Q, q] \in N : [Q, q] \mathrel{\overline{\subset}} [H, h]\}$. $\square$

Finally, note that the multigraph constructed in the proof of Lemma 84 still satisfies the updated definition of multigraph and witnesses the truth of the axiom of Infinity in this new model. Therefore we have proved:

Theorem 8. *It is consistent with the rest of our multiset theory that the inclusion relation is not antisymmetric.*

4 Index

4.1 Notations

4.2 Definitions

References

[Aczel 1] Aczel, P. Non well-founded sets. 1988. *CSLI - Lecture Notes, vol 14, Stanford University, Center for the Study of Language and Information.*

[Blizard 1] Blizard, W. D. The development of multiset theory. 1991. *Modern Logic, Volume 1, Number 4.*

[Blizard 2] Blizard, W. D. Multiset theory. 1989. *Notre Dame Journal of Formal Logic, Volume 30, Number 1.*

[Blizard 3] Blizard, W. D. Dedekind multisets and function shells. 1993. *Theoretical Computer Science, Volume 110, Issue 1.*

[Blizard 4] Blizard, W. D. Negative membership. 1990. *Notre Dame Journal of Formal Logic, Volume 31, Number 3.*

[Coret 1] Coret, J. Sur les cas stratifiés du schema de remplacement. 1970. *Comptes Rendus hebdomadaires des séances de l'Académie des Sciences de Paris série A.*

[Dang and McKenzie 1] Dang, V. and McKenzie, Z. A permutation method yielding models of the stratified axioms of Zermelo Fraenkel set theory. 2009. *Cahiers du Centre de logique, Volume 16, Academia-Bruylant, Louvain-la-Neuve (Belgium).*

[Ehrenfeucht 1] Ehrenfeucht, A. An application of games to the completeness problem for formalized theories. 1960. *Fundamenta Mathematicae, Volume 49.*

[Felgner 1] Felgner, U. Models of ZF-set theory. 1971. *Lecture Notes in Mathematics Volume 223, Springer, Berlin.*

[Forster 1] Forster, T. E. AC fails in the natural analogues of V and L that model the stratified fragment of ZF. 2003. *Contemporary Mathematics, vol 361, American Mathematical Society.*

[Hinnion 1] Hinnion, R. Sur la théorie des ensembles de Quine. 1975. *PhD Thesis. ULB, Brussels.*

[Holmes 1] Holmes, M. R. Elementary Set Theory with a Universal Set. 1998. *Cahiers du Centre de logique, Volume 10, Academia, Louvain-la-Neuve (Belgium).*

[Mathias 1] Mathias, A. R. D. Weak systems of Gandy, Jensen and Devlin. 2006. *Set Theory: Centre de Recerca Matemàtica, Barcelona 2003-4, edited by Joan Bagaria and Stevo Todorcevic, Trends in Mathematics, Birkhäuser Verlag, Basel.*

[Tarski 1] Tarski, A. A lattice-theoretical fixpoint theorem and its applications. 1955. *Pacific Journal of Mathematics, Volume 5, Number 2.*