Universiteit Utrecht

Department of Mathematics

Bachelor thesis

Ordinals, Cardinals
& Numerosities

Author:
Eric Faber (3365026)

Supervisor:
Dr. Jaap van Oosten

July 22, 2011

Abstract

The concept of a numerosity of a set tends to provide a natural notion of size on infinite sets. For example, it should preserve the ordering by strict inclusion. Cardinal numbers do certainly not satisfy this property.

By defining a numerosity function as a function on sets that preserves this ordering, and having values, called the numerosities, endowed with the structure of a semi-ring, one can prove that the existence of such a function depends on certain types of ultrafilters.

A numerosity function can be defined on a class of countable labelled sets provided a Ramsey ultrafilter on the set of natural numbers exists. Furthermore, this function can be exploited to prove the existence of a nonstandard universe, which provides applications for the notion of a numerosity in nonstandard analysis.

On arbitrary sets of ordinals, existence of a less constrained numerosity function can be proved as well, albeit much harder and less elegant than in the countable case.

Contents

Chapter 1

Introduction

In set theory, mathematicians have always been interested in finding a good measure of the size of sets. Probably the reader is already familiar with the concept of cardinality, which gives a notion of size based on one-to-one correspondence between sets. However, the behavior of infinite sets with respect to cardinality is far from intuitive. The challenge is to extend the idea of measuring a finite set to infinite sets as naturally as possible, i.e. with as many "finite" properties as possible.

For instance, a basic requirement would be that proper subsets of a set are strictly smaller than their majorant. Adopting the terminology of [2], this will be called the *Aristotelian* property, referring to Aristotle's principle "the whole is more than the sum of its parts". Obviously cardinality is *non-aristotelian.* However, cardinality has the *Cantor* property of measuring the size of sets, namely that two sets between which there is a one-to-one correspondence, are equally large.

The two properties described above are widely known to be incompatible. But omitting the *Cantor* property and imposing extra arithmetical conditions give rise to the notion of a *numerosity* [1], which seems a natural extension of "counting" finite sets.

Chapter 2 is devoted to the notions of *ordinal* and *cardinal* numbers. These notions give an insight in what kind of behaviour is to expect from infinite sets. Mostly this is far from a "finite" behaviour, but there are interesting results.

Then chapter 3 starts by assuming a *numerosity* function on the class of countable *labelled sets.* It is defined as an *Aristotelian* notion of size. In the case of countable sets, this definition leads in a very natural way to a proof that the set of numerosities is a set of hyperintegers, so it is basically a set of integers, but it contains infinite elements. So a notion of size that behaves like a finite notion of size, leads to infinite integers that behave like finite ones. A numerosity function can then be used to define a nonstandard universe, a mathematical construction that justifies the theory of nonstandard analysis, which is a very interesting consequence.

At last, chapter 4 makes an attempt to extend this notion of size to arbitrary sets of ordinals. The result is here basically the same, although the proofs are less elegant and insightful. However, this chapter shows many of the unusual properties of ordinals, and poses conditions that make set operations on ordinals easier.

Overall, this thesis provides an introduction in the theory of ordinal and cardinal numbers, and shows their mathematical importance by explaining the new concept of a numerosity. Some concepts are explained in the appendix, the assumed knowledge is only a little model theory and basic set theory.

Chapter 2

Ordinal and Cardinal numbers

In this chapter two main concepts for measuring the size of sets are briefly introduced, as the idea of a *numerosity* makes use of these concepts. Some techniques in the proofs may be used later on, so this chapter forms a strong basis for the main subject. The parts on ordinal and cardinal arithmetic are predominantly used in chapter 4.

Most of the proofs in this chapter are adopted from [5] and [4], and slightly extended.

2.1 Ordinals

Ordinals can be most intuitively viewed as the concept of counting elements extended to the infinite. It is founded in the idea of a *transitive* set, which will be defined first.

Definition 2.1.1. *A set A is called* ***transitive*** *if for all $a \in A$, we have $a \subset A$.*

An interesting property of transitive sets is the following:

Proposition 2.1.2. *Let a be a transitive set. Then for $A = a \cup \{a\}$, A is transitive.*

Proof. Let $x \in A$ be arbitrary. If $x \in a$, then $x \subset a$ (a is transitive), hence $x \subset A$. If $x = a$, obviously $x \subset A$ □

With these notions in mind, an *ordinal number* can be defined:

Definition 2.1.3. *An* ***ordinal****, or* ***ordinal number*** *is a transitive set well-ordered by the (strict) order "$\in$".*

The smallest transitive set is the empty set $\varnothing$, which is by definition also an ordinal (the smallest ordinal, as will become clear). From proposition 2.1.2 and the definition of an ordinal follows that for α an ordinal, $\alpha \cup \{\alpha\}$ is an ordinal as well. This leads to the first arithmetical property of ordinals, the existence of a successor.

Definition 2.1.4. *For α an ordinal, the* ***successor*** *of α, $S(\alpha)$ or $\alpha + 1$ is defined as: $S(\alpha) = \alpha \cup \{\alpha\}$*

From the ordinal $\varnothing$, the finite ordinals $S(\varnothing), S(S(\varnothing)), S(S(S(\varnothing)), \ldots$ or:

$$\{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}, \ldots$$

are in fact a set-theoretical foundation of the natural numbers $1, 2, 3, \ldots$. The relation of elementhood "$\in$" can also be used to compare ordinals themselves. With this relation, the successor $S(\alpha)$ is the smallest ordinal bigger than α, so the name "successor" is appropriate.

Notice each element of these finite ordinals is an ordinal itself. This holds in general: Let α be an ordinal, and $\beta \in \alpha$. Suppose $\gamma \in \beta$. Then $\delta \in \gamma$ implies $\delta \in \beta$, as α is well-ordered by $\in$. So $\gamma \subset \beta$, therefore β is transitive and well-ordered by the restricted order, hence β is an ordinal. In fact, any ordinal contains precisely all smaller ordinals.

The following proposition states more basic properties of ordinals:

Proposition 2.1.5. *(i) For $\alpha \subsetneq \beta$ both ordinals, $\alpha \in \beta$ (ii) For $\alpha \neq \beta$ ordinals, $\alpha \in \beta$ or $\beta \in \alpha$.*

Proof. (i) β is an ordinal, so let γ be the smallest element of $\beta \setminus \alpha$. Then for any $\delta \in \gamma$, we must have $\delta \in \alpha$, hence $\gamma \subseteq \alpha$. α is an ordinal, so any $\delta \in \alpha$ is one as well. As δ and γ are both in β, We have $\gamma \in \delta$ or $\gamma = \delta$ or $\delta \in \gamma$. By definition $\gamma \notin \alpha$, so the first two options cannot happen, because then $\gamma \in \delta \subset \alpha$ or $\gamma = \delta \in \alpha$. Hence $\delta \in \gamma$, so $\alpha \subseteq \gamma$, hence it follows $\alpha = \gamma \in \beta$.

(ii) Let $\gamma = \alpha \cap \beta$. Then γ is an ordinal. Suppose $\gamma \neq \alpha$ and $\gamma \neq \beta$. Then, by (i), $\gamma \in \alpha$ and $\gamma \in \beta$ hence $\gamma \in \gamma$, a contradiction. So $\gamma = \alpha$ or $\gamma = \beta$, hence $\alpha \subset \beta$ or $\beta \subset \alpha$. Applying (i) once more finishes the proof. ☐

The following corollary is a very nice property of ordinals, and will be very useful:

Corollary 2.1.6. *For $\mathcal{W}$ a set of ordinals, a supremum and infimum is defined: $\sigma = \bigcup \mathcal{W} = \sup \mathcal{W}$, $\iota = \bigcap \mathcal{W} = \inf \mathcal{W}$.*

Proof. By the previous proposition, σ and ι are both transitive, and each element is an ordinal. Furthermore, if for an ordinal α, $\alpha \notin \sigma$, then forall $\beta \in \mathcal{W} : \alpha \notin \beta$, so then $\alpha = \sup \mathcal{W}$ or $\sup \mathcal{W} \in \alpha$. Hence σ contains precisely all ordinals smaller than $\sup \mathcal{W}$. It follows $\sigma = \sup \mathcal{W}$.

For ι, similar considerations prove $\iota = \inf \mathcal{W}$. ☐

An important theorem about ordinals is the following:

Theorem 2.1.7. *Every well-ordered set $< W, R >$ is isomorphic to a unique ordinal number.*

Proof. For $W(x) = \{y \in W | yRx\}$ an initial segment of W, define $F(x) = \alpha$ if $W(x) \cong \alpha$ for an ordinal α. If α exists, it is unique, as no well-order can be isomorphic to an initial segment of itself, especially ordinals cannot (composition of the isomorphisms would yield this result). For each $x \in W$, such $F(x)$ exists, otherwise consider the least element x_0 such that $W(x_0)$ is not isomorphic to an ordinal. Then $\alpha = \cup_{xRx_0, x \neq x_0} F(x)$ is the least ordinal bigger than $F(x)$ for all xRx_0, so $F(x_0) = \alpha$ would then extend F appropriately to an isomorphism $F : W(x_0) \to \alpha$, a contradiction. So $F(x)$ exists for all $x \in W$.

Let now be $\gamma = \cup_x F(x) = \sup_x F(x)$. Then $W \cong \gamma$, the isomorphism is given by F. □

This theorem presents us with a notion of size for well-ordered sets. For a well-order $\langle W, R\rangle$, the corresponding ordinal according to the previous theorem is called the *order type* of the well-order, or type$\langle W, R\rangle$.

It is important to remark that the class *Ord* of all ordinals is not a set, otherwise sup *Ord* would be a proper ordinal, and its successor would not be in *Ord*. By proposition 2.1.5, *Ord* is linearly ordered by "$\in$". In the following, for ordinals α, β, by $\alpha < \beta$ is meant $\alpha \in \beta$, and $\alpha \leq \beta$ is the relaxed version of this, so in that case it means $\alpha \in \beta$ or $\alpha = \beta$

Ordinals have arithmetical properties, to which the next section is devoted.

2.1.1 Ordinal arithmetic

A *limit ordinal* is an ordinal that is not a successor of another ordinal. For example, let $\mathcal{W}_{\text{fin}}$ be the set of finite ordinals, then $\omega = \sup \mathcal{W}_{\text{fin}} = \cup\mathcal{W}_{\text{fin}}$ is a limit ordinal, in this case it is the first infinite ordinal, isomorphic to $\langle \mathbb{N}, \leq\rangle$. With $n \in \omega$ is often meant $n \in \mathbb{N}$. Finite ordinals satisfy the same arithmetical properties as natural numbers (as will become clear), so there's no harm in this ambiguity.

Transfinite induction is an important tool in proving statements about ordinals, therefore it will be stated here in terms of ordinals:

Definition 2.1.8. *For $\mathcal{W}$ a class of ordinals, suppose the following holds:*

(i) $0 \in \mathcal{W}$

(ii) $\mathcal{W}$ is closed under the successor function S

(iii) for α a limit ordinal, if $\beta < \alpha$ implies $\beta \in \mathcal{W}$, then $\alpha \in \mathcal{W}$.

Then $\mathcal{W}$ =Ord.

For nondecreasing sequences $\langle a_\xi, \xi < \beta\rangle$, there is a notion of a limit:

Definition 2.1.9. *Let $\langle a_\xi, \xi < \beta\rangle$ be a nondecreasing sequence, then:*

$$\lim_{\xi \to \theta} a_\xi = \sup_{\xi < \theta} a_\xi$$

With these tools, addition, multiplication and exponentiation of ordinals can be defined inductively. In the following definitions, the parts (i), (ii) and (iii) refer to the corresponding parts of definition 2.1.8 that is used to define the operation.

Definition 2.1.10 (Addition). *(i)* $\alpha + 0 = \alpha$

(ii) $\alpha + S(\beta) = S(\alpha + \beta) = \alpha + \beta + 1$

(iii) For $\beta \neq 0$ a limit ordinal, $\alpha + \beta = \lim_{\xi \to \beta} \alpha + \xi$

With this definition, it follows immediately that addition is not commutative:

$$1 + \omega = \lim_{\xi \to \omega}(1 + \xi) = \omega \neq \omega + 1 = \omega \cup \{\omega\}$$

Definition 2.1.11 (Multiplication)**.** *(i)* $\alpha \cdot 0 = 0$

(ii) $\alpha \cdot S(\beta) = \alpha \cdot (\beta + 1) = \alpha \cdot \beta + \alpha$

(iii) For $\beta \neq 0$ *a limit ordinal,* $\alpha \cdot \beta = \lim_{\xi \to \beta} \alpha \cdot \xi$

Multiplication is also noncommutative:

$$2 \cdot \omega = \lim_{\xi \to \omega}(2 \cdot \xi) = \omega \neq \omega \cdot 2 = \omega \cdot (1 + 1) = \omega + \omega = \lim_{\xi \to \omega}(\omega + \xi)$$

Exponentiation is defined in the following manner:

Definition 2.1.12 (Exponentiation)**.** *(i)* $\alpha^0 = 1$

(ii) $\alpha^{S(\beta)} = \alpha^{\beta+1} = \alpha^\beta \cdot \alpha$

(iii) For $\beta \neq 0$ *a limit ordinal,* $\alpha^\beta = \lim_{\xi \to \beta} \alpha^\xi$

The following may confuse the less experienced reader:

$$2^\omega = \lim_{\xi \to \omega} 2^\xi = \sup_{n < \mathbb{N}} 2^n = \omega$$

The *ordinal* number 2^ω must not be confused with the *cardinal* number $2^\omega = \aleph_1 \neq \aleph_0 = \omega$. In fact, $\omega^\omega = \aleph_1$ is the first uncountable ordinal. In the next section there will be more about cardinal numbers.

There is also an explicit (i.e. non-recursive) statement for addition and multiplication, that will be needed later on. The definition is equivalent, which will not be proven here, but it is done by induction.

Definition 2.1.13 (Addition, Multiplication)**.**

(i) $\alpha + \beta = \mathit{type}\langle \alpha \times \{0\} \cup \beta \times \{1\}, R\rangle$ *where R is defined as follows:*

$$\begin{aligned} (\xi, 0)R(\eta, 0) &\iff \xi < \eta < \alpha \\ (\xi, 1)R(\eta, 1) &\iff \xi < \eta < \beta \\ (\xi, 0)R(\eta, 1) &\iff \xi < \alpha, \eta < \beta \end{aligned}$$

(ii) $\alpha \cdot \beta = \mathit{type}\langle \beta \times \alpha, R\rangle$ *where R is defined by lexicographical order on* $\beta \times \alpha$*:*

$$(\xi, \eta)R(\xi', \eta') \iff (\xi < \xi') \vee ((\xi = \xi') \wedge (\eta < \eta'))$$

There are some more arithmetical properties of ordinal numbers, that are actually very similar to those of integers. Associativity of addition and multiplication follows by induction in a trivial way. In the same way it follows that addition, multiplication of ordinals > 0, exponentiation of ordinals > 1 are all order-preserving. The following lemma will be useful in proving an interesting theorem on ordinals:

Lemma 2.1.14 (Difference and division of ordinals)**.** *(i) For* α, β *ordinals, with* $\alpha < \beta$*, there exists a unique* $\delta \in \mathit{Ord}$ *such that* $\alpha + \delta = \beta$*.*

(ii) *For α, β arbitrary ordinals with $\alpha > 0$, there exist unique γ, ρ such that $\rho < \alpha$ and $\beta = \alpha \cdot \gamma + \rho$.*

Proof. (i) Let $\delta = \text{type}\langle \xi : \alpha \leq \xi < \beta, < \rangle$. Notice $\beta = \text{type}\langle \xi : \xi < \beta, < \rangle$, and $\alpha = \text{type}\langle \xi : \xi < \alpha, < \rangle$. According to definition 2.1.13 it is easily seen that $\beta = \alpha + \delta$.

(ii) let $\gamma = \sup\{\epsilon : \alpha \cdot \epsilon < \beta\}$. By (i), it follows $\alpha \cdot \gamma \leq \beta$, and again by (i) there exists ρ such that $\beta = \alpha \cdot \gamma + \rho$. If $\rho \geq \alpha$, then $\alpha \cdot (\gamma + 1) \leq \beta$, which contradicts the choice of γ. So $\rho < \alpha$.

□

The important result is now the following theorem:

Theorem 2.1.15 (Cantor's normal form)**.** *For every ordinal α there exists ordinals $\alpha \geq \alpha_1 > \alpha_2 > \cdots > \alpha_n$, and natural numbers (i.e. finite ordinals) $h_1, \ldots, h_n$ such that:*

$$\alpha = \omega^{\alpha_1} h_1 + \cdots + \omega^{\alpha_n} h_n$$

*This representation, called the **normal form** of α, is unique.*

Proof. The proof is by induction on α. For $\alpha = 1$, obviously $\alpha = \omega^0 \cdot 1$. Trivially, the set of ordinals for which such a unique representation exists is closed under succession. Suppose α is a limit element, and the theorem holds for all $\beta < \alpha$. Let α_1 be the greatest ordinal such that $\omega^{\alpha_1} \leq \alpha$. Then, by lemma 2.1.14, there exists γ, ρ such that $\alpha = \omega^{\alpha_1} \cdot \gamma + \rho$ and $\rho < \omega^{\alpha_1}$. Observe γ must be finite, or else $\omega \leq \gamma$, so then $\omega^{\alpha_1 + 1} = \omega^{\alpha_1}\gamma \leq \alpha$ which contradicts the choice of α_1. By hypothesis, a unique representation for ρ exists, and by the division lemma the found representation for α is unique as well. □

The previous theorem makes arithmetic with ordinals a lot easier. An interesting phenomenon is that of absorption, examples of this have been discussed earlier:

$$2 \cdot \omega = \omega$$
$$2 + \omega = \omega$$

The 2 is "absorbed" by the much larger ordinal "ω". Because "ω" is the first infinite ordinal, it is easy to see when absorption occurs in these examples. By Cantor's normal form theorem, these examples can be generalized to all ordinals in a simple way. Suppose that $\alpha = \omega^{\alpha_1} h_1 + \cdots + \omega^{\alpha_n} h_n$, $\beta = \omega^{\beta_1} k_1 + \cdots + \omega^{\beta_m} k_m$. Then:

$$\alpha \cdot \beta = \sum_{i=1}^{n} \sum_{j=1}^{m} \omega^{\alpha_i} h_i \omega^{\beta_j} k_j$$

Notice that $\beta_j > 0$ for $1 \leq j < m$, so absorption of the integers h_i occurs:

$$\begin{aligned} \alpha \cdot \beta &= \textstyle\sum_{i=1}^{n} \sum_{j=1}^{m-1} \omega^{\alpha_i} \omega^{\beta_j} k_j + \sum_{i=1}^{n} \omega^{\alpha_i} h_i \omega^{\beta_m} k_m \\ &= \textstyle\sum_{i=1}^{n} \sum_{j=1}^{m-1} \omega^{\alpha_i + \beta_j} k_j + \sum_{i=1}^{n} \omega^{\alpha_i} h_i \omega^{\beta_m} k_m \end{aligned}$$

Now the exponent terms $\alpha_i + \beta_j$ might be reduced as well, and in this way the product $\alpha \cdot \beta$ can be written in normal form in a recursive manner. This method is useful when looking for ordinals with certain absorption properties. For example, what will be useful later on is an ordinal θ that satisfies the property that $\alpha \cdot \beta + \gamma < \theta$, whenever $\alpha, \beta, \gamma < \theta$. Such ordinals will be called *arithmetically closed*, or *atoms*, as in [2].

Such an ordinal should be necessarily of the form $\theta = \omega^{\xi}$, for ξ an ordinal ≥ 0 (otherwise θ is a sum of two smaller terms already!). Furthermore, observe that for the highest exponents α_1, β_1 of the normal forms of α, β, one must have:

$$\alpha_1 + \beta_1 < \xi$$

Because θ is of the form ω^{ξ}, it is clear that $\alpha_1, \beta_1 < \xi$ for $\alpha, \beta < \theta$. For the above to hold, notice ξ should be of the form ω^{ζ}, for $\zeta \geq 0$, by the exact same argument as before. In fact, the smallest atom is given by $\theta = \omega^{\omega^0} = \omega$. Indeed, for every $\alpha, \beta, \gamma < \omega$, $\alpha \cdot \beta + \gamma < \omega$.

In general, the above shows that every atom is of the form $\theta = \omega^{\omega^{\zeta}}$, for $\zeta \geq 0$.

2.2 Cardinals

Cardinality is one of the concepts Georg Cantor, the inventor of set theory, is most famous for. This section is mostly devoted to rephrase some important concepts for use later on. Most of the ideas are assumed familiar to the reader though, so the proofs, if stated, will be concise.

In the first place, cardinality is thought of as an equivalence relation on sets. For A, B sets, we say $|A| = |B|$ if and only if there exists a bijective function $f : A \to B$. Among the widely known examples of this relation are $|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}| \neq |\mathbb{R}|$. One writes $|A| \leq |B|$ if there is an injective map $f : A \to B$, and $|A| < |B|$ if $|A| \leq |B|$ but not $|A| = |B|$. One of the most important theorems involving this relation is the following, which will be stated without proof:

Theorem 2.2.1 (Schröder-Cantor-Bernstein)**.** *Let A, B be sets. If $|A| \leq |B|$ and $|B| \geq |A|$, then $|A| = |B|$.*

A definition of $|A|$ itself can be given in terms of ordinals.

Definition 2.2.2. *The cardinality $|A|$ is the least ordinal α such that $|\alpha| = |A|$. Hence assuming AC (the Axiom of Choice), $|A|$ is defined for every A, as any well-order is isomorphic to a unique ordinal number (by theorem 2.1.7), so there is at least one such bijection. An ordinal number α is a **cardinal number** if and only if $|\alpha| = \alpha$.*

All finite ordinals are cardinals, as is the least infinite ordinal number $\aleph_0 = \omega$. For α a cardinal, a successor cardinal α^+ is defined as the least cardinal $> \alpha$. A cardinal κ is a limit cardinal if it is not a successor cardinal, so there is no α such that $\kappa = \alpha^+$. A general notion is the *aleph number*:

Definition 2.2.3 (Aleph number)**.** *The **aleph numbers** $\aleph_\alpha$ are defined inductively as:*

(i) $\aleph_0 = \omega$

(ii) $\aleph_{\alpha+1} = {\aleph_\alpha}^+$

(iii) For κ a limit ordinal, $\aleph_\kappa = \sup \{\aleph_\alpha | \alpha < \kappa\}$

It is easy to check that the class of aleph numbers is precisely the class of infinite cardinals (assuming one can well-order any set), ordered by their ordinal index.

As is well known,$|2^X| = |\mathscr{P}(X)| > |X|$, for any set X; this can be shown by a diagonal argument. For the aleph numbers, we therefore have $2^{\aleph_\alpha} = |\mathscr{P}(\aleph_\alpha)| \geq \aleph_{\alpha+1}$ (recall the ambiguity in notation). Cantor initially speculated $2^{\aleph_0} = \aleph_1$. This is known as the *continuum hypothesis*, shortened as *CH*. The conjecture that $2^{\aleph_\alpha} = \aleph_{\alpha+1}$ for all α is known as the *generalized continuum hypothesis*, or *GCH*.

Cardinal arithmetic should be distinguished from ordinal arithmetic. I'll use the same symbols for multiplication and addition, as it should be clear from the context what kind of multiplication is meant. The same holds for exponentiation.

Definition 2.2.4. *Let κ, λ be cardinals. Define addition, multiplication and exponentiation on cardinals as:*

$$\begin{aligned}&\kappa + \lambda = |\{\kappa \times \{0\} \cup \lambda \times \{1\}\}|\\ &\kappa \cdot \lambda = |\kappa \times \lambda|\\ &\kappa^\lambda = |{}^\lambda\kappa| = |\{f \in \mathscr{P}(\kappa \times \lambda) : f : \lambda \to \kappa \text{ a function}\}|\end{aligned}$$

It follows immediately from definition 2.1.13 that both + and $\cdot$ are commutative. Note that for any finite n, a function in ${}^n\kappa$ is just a tuple $(\mu_1, \dots, \mu_n)$ with $\mu_i \in \kappa$. So $\kappa^n = \kappa \cdot \kappa \cdots \kappa$ (n times).

With the newly defined notation, notice $\aleph_0 \cdot \aleph_0 = \aleph_0$, a result that should be familiar. This result can be generalized, and what is needed is a well-ordering $\lhd$ on the class *Ord* × *Ord*. Define $\lhd$ on *Ord* × *Ord* as follows:

$$(\alpha, \beta) \lhd (\gamma, \delta) \iff \begin{array}{l}\max\{\alpha, \beta\} < \max\{\gamma, \delta\} \text{ or}\\ \max\{\alpha, \beta\} = \max\{\gamma, \delta\} \text{ and } \alpha < \gamma \text{ or}\\ \max\{\alpha, \beta\} = \max\{\gamma, \delta\}\ ,\ \alpha = \gamma \text{ and } \beta < \delta\end{array}$$

For any α, notice that for the initial segment defined by $(0, \alpha)$ holds:

$$\{(\gamma, \delta) \in \mathit{Ord} \times \mathit{Ord} : (\gamma, \delta) \lhd (0, \alpha)\} = \alpha \times \alpha.$$

Also, for any α, β, it holds for the initial segment defined by (α, β):

$$\{(\gamma, \delta) \in \mathit{Ord} \times \mathit{Ord} : (\gamma, \delta) \lhd (\alpha, \beta)\} \subseteq (\max\{\alpha, \beta\} + 1) \times (\max\{\alpha, \beta\} + 1).$$

Consider the following lemma:

Lemma 2.2.5. *For κ an infinite cardinal, $\kappa \cdot \kappa = \kappa$.*

Proof. Notice this holds for the smallest infinite cardinal $\omega = \aleph_0$. The rest of the proof is by induction. Let $\lhd$ be the well-ordering on *Ord* × *Ord* defined above. For any $\alpha < \kappa$ $|\alpha \times \alpha| = |\alpha| \cdot |\alpha| < \kappa$ by induction hypothesis, and also for finite α. With this, and the above, we can estimate the cardinality of the initial segment of any $(\alpha, \beta) \in \kappa \times \kappa$, ordered by $\lhd$:

$$|\{(\gamma, \delta) : (\gamma, \delta) \lhd (\alpha, \beta)\}| \leq |(\max\{\alpha, \beta\} + 1) \times (\max\{\alpha, \beta\} + 1)| < \kappa$$

The latter follows because any infinite cardinal is a limit ordinal: $|\lambda + 1| = |1 + \lambda| = |\lambda|$ for any ordinal λ. So type$\langle \kappa \times \kappa \rangle \leq \kappa$. Therefore $|\kappa \times \kappa| \leq \kappa$ (there exists an injection $i : \kappa \times \kappa \to \kappa$). So $\kappa \cdot \kappa = \kappa$ for all infinite κ, by induction. □

An important consequence is the following theorem, which will be the most important result here.

Theorem 2.2.6. *Let* $\aleph_\alpha, \aleph_\beta$ *be aleph numbers. Then:*

$$\aleph_\alpha + \aleph_\beta = \aleph_\alpha \cdot \aleph_\beta = \max\{\aleph_\alpha, \aleph_\beta\}$$

Proof. The product follows from:

$$\max\{\aleph_\alpha, \aleph_\beta\} \leq \aleph_\alpha \cdot \aleph_\beta \leq \max\{\aleph_\alpha, \aleph_\beta\} \cdot \max\{\aleph_\alpha, \aleph_\beta\} = \max\{\aleph_\alpha, \aleph_\beta\}$$

The sum follows from:

$$\max\{\aleph_\alpha, \aleph_\beta\} \leq \aleph_\alpha + \aleph_\beta \leq \aleph_\alpha \cdot \aleph_\beta$$

□

Under AC any infinite cardinal is an aleph, so the above holds for any two infinite cardinals.

A cardinal κ is called *regular* if for any ordinal $\mu < \kappa$, and every map $f : \mu \to \kappa$ we have $\sup\{f(\nu) | \nu \in \mu\} < \kappa$.

For example, the ordinal $\gimel = \aleph_{\aleph_{\aleph_{(\omega \text{ times})}}}$ is not regular: define $f : \omega \to \gimel$ as $f(i) = \aleph_{\aleph_{\ldots \aleph_0 (i \text{ times})}}$, then $\sup\{f(\nu) | \nu \in \omega\} = \gimel$. The cardinal $\aleph_0$ is regular, as any finite sequence has a finite upper bound. Also $\aleph_1$ is regular, for it is the first uncountable cardinal. Indeed, let $\mu < \aleph_1$ be arbitrary, and $f : \mu \to \aleph_1$ an arbitrary function, then:

$$\sup\{f(\nu) | \nu \in \mu\} = \bigcup \{f(\nu) | \nu \in \mu\}$$

The latter yields a countable union of countable sets, so the result is still countable.

An important remark is that the latter only holds by assuming the axiom of choice. This holds for many properties of cardinal numbers, especially arithmetical properties. The following lemma has as consequence that any $\aleph_{\alpha+1}$ is regular:

Lemma 2.2.7. *For* S *a set,* $|\bigcup S| \leq |S| \cdot \sup\{|X| : X \in S\}$.

Proof. Let $\kappa = |S|$ and $\mu = \sup\{|X| : X \in S\}$. Notice μ is a cardinal, for else if there is a bijection $f : \mu \to \alpha$ for some $\alpha < \mu$, then there is $X \in S$ such that $\alpha < |X| \leq \mu$, so then the restriction $f : |X| \to \alpha$ gives $|X| \leq \alpha$ which is a contradiction.

There is a bijection between κ and S, so the elements of S can be written as X_α for $\alpha < \kappa$, for unique α. For any such X_α, let $\mu_\alpha = |X_\alpha| \leq \mu$. By the same argument, the elements of X_α can be denoted as $x_{\alpha,\beta}$ for $\beta < \mu_\alpha$. Note AC is used in this step, as for any X_α one has to choose such a representation. Thus the map $x_{\alpha,\beta} \mapsto (\alpha, \beta)$, for $\alpha < \kappa, \beta < \mu$ is an injection $\bigcup S \to \kappa \times \mu$, hence:

$$|\bigcup S| \leq |\kappa \times \mu| = \kappa \cdot \mu = |S| \cdot \sup\{|X| : X \in S\}$$

□

If any successor aleph $\aleph_{\alpha+1}$ is not regular, then it is the union of at most $\aleph_\alpha$ sets with cardinality $\aleph_\alpha$ (taking the supremum as union). From the lemma, it follows that then $\aleph_{\alpha+1} \leq \aleph_\alpha \cdot \aleph_\alpha = \aleph_\alpha$ which is contradictory.

Chapter 3

Numerosities

This chapter introduces a the notion of a *numerosity* of a labelled set. First only countable sets are considered, for uncountable sets the situation is more involved. That case will be treated in chapter 4. Most definitions and theorems are given as in [1], the proofs are extended versions of the proofs stated there.

3.1 From finite approximations to numerosities

For countable sets, the key idea of assigning numerosities to sets is to approximate this set by finite sets. This is done by first *labelling* the set appropriately:

Definition 3.1.1. *A labelled set* $\mathbf{A}$ *is a set* A *together with a map* $\ell_A : A \to \mathbb{N}$ *such that for all* $i \in \mathbb{N}$, $|\ell_A^{-1}(i)| < \omega$. *The set* A *is called the* ***domain*** *of* $\mathbf{A}$. *The class of labelled sets is denoted by* $\mathscr{L}$.

In analogy with [1], a labelled set is printed in bold face, while its domain is denoted by the same symbol in normal face. For a labelled set $\mathbf{A}$, a *finite approximation* is given by the sequence $A_i = \{a \in A | \ell_A(a) \le i\}$. That is, a labelled set $\mathbf{A}$ can be "approximated" by the sets $A_0 \subseteq A_1 \subseteq \ldots \subseteq A_i \subseteq \ldots$:

$$A = \bigcup_{i=0}^{\infty} A_i$$

The finite cardinalities $|A_i|$ of these sets can be thought of as an approximating sequence of the actual numerosity $\mathfrak{n}(A)$ of the set A.

Now the class of labelled sets can be provided with some arithmetic on labelled sets:

Definition 3.1.2. $\mathbf{A} \oplus \mathbf{B}$ *is a labelled set defined by the disjoint sum* $A \uplus B$ *and the labelling function* $\ell_{A \uplus B}$, *defined as:*

$$\ell_{A \uplus B}(x) = \begin{cases} \ell_A(x) & \text{if } x \in A \\ \ell_B(x) & \text{if } x \in B \end{cases}$$

$\mathbf{A} \odot \mathbf{B}$ *is a labelled set defined by the cartesian product* $A \times B$ *and the labelling function* $\ell_{A \times B}$:

$$\ell_{A \times B}(x, y) = max\{\ell_A(x), \ell_B(y)\}$$

Note that $|(A \oplus B)_i| = |A_i| + |B_i|$ and $|(A \odot B)_i| = |A_i| \cdot |B_i|$. A numerosity function on labelled sets is now defined as follows:

Definition 3.1.3. *A* ***numerosity function*** *is a surjective map* $\mathfrak{n} : \mathscr{L} \twoheadrightarrow \mathscr{N}$, *where* $\mathscr{N}$ *is a linearly ordered set of* ***numerosities****. It has the following properties:*

(i) $\forall i \in \mathbb{N} : |A_i| \leq |B_i|$ *implies* $\mathfrak{n}(\mathbf{A}) \leq \mathfrak{n}(\mathbf{B})$.

(ii) The difference principle: $\xi < \mathfrak{n}(\mathbf{A})$ *if and only if* $\xi = \mathfrak{n}(\mathbf{B})$ *for some* $\mathbf{B} \subset \mathbf{A}$. *With the latter is meant* $B \subset A$ *and* $\ell_B = \ell_A|_B$.

(iii) For $\mathfrak{n}(\mathbf{A}) = \mathfrak{n}(\mathbf{A}')$ *and* $\mathfrak{n}(\mathbf{B}) = \mathfrak{n}(\mathbf{B}')$ *holds: The sum principle:* $\mathfrak{n}(\mathbf{A} \oplus \mathbf{B}) = \mathfrak{n}(\mathbf{A}' \oplus \mathbf{B}')$ *and the cartesian product principle:* $\mathfrak{n}(\mathbf{A} \odot \mathbf{B}) = \mathfrak{n}(\mathbf{A}' \odot \mathbf{B}')$

The aristotelian property, as explained in the introduction, is satisfied here, it follows directly from the "if" part of property (ii). That is precisely the part that the notion of cardinality does not satisfy.

The set $\mathscr{N}$ remains a bit vague here. It will be proven later that it can be seen as a set of "hypernatural numbers", which is basically a model of natural numbers containing non-standard elements. This is one of the implications of the very "natural" properties that are postulated in the definition of $\mathfrak{n}$.

Another thing that is important to keep in mind is that the numerosity of a set A depends on the chosen labelling ℓ_A. There's a need for some sort of "canonical" labelling for all sets, which is another discussion. For subsets of $\mathbb{N}$, however, a canonical labelling is obvious, just let this be the identity function.

Some basis properties of numerosities become clear in the following proposition:

Proposition 3.1.4. *(i)* $\mathscr{N}$ *has a least element that we call* $0 = \mathfrak{n}(\varnothing)$, *for* $\varnothing$ *the empty labelled set.*

(ii) All labelled singletons have numerosity 1.

(iii) For any numerosity, we have a successor: for $\xi = \mathfrak{n}(A)$, *let* $\xi + 1 = \mathfrak{n}(\mathbf{A} \oplus \{\mathbf{y}\})$, *where* $\{\mathbf{y}\} \in \mathscr{L}$ *is an arbitrary (labelled) singleton. Also, for* $\mathbf{A} \neq \varnothing$, $\mathfrak{n}(\mathbf{A})$ *has a predecessor.*

(iv) For $\mathbf{A}$ *a finite labelled set,* $\mathfrak{n}(\mathbf{A}) = |A|$

Proof.

(i) This follows from the difference principle, as there are no proper subsets of the empty set

(ii) Suppose this is not the case. Then there are two labelled singletons $\{\mathbf{y}\}$ and $\{\mathbf{z}\}$, such that $\mathfrak{n}(\{\mathbf{z}\}) < \mathfrak{n}(\{\mathbf{y}\})$. Then, again by the difference principle, there is a proper subset $\mathbf{A} \subset \{\mathbf{y}\}$ such that $\mathfrak{n}(\mathbf{A}) = \mathfrak{n}(\{\mathbf{z}\})$. But the only proper subset of a singleton is the empty set, hence $\mathfrak{n}(\mathbf{A}) = \mathfrak{n}(\{\mathbf{z}\}) = 0$, which is a contradiction (we must have $\mathfrak{n}(\varnothing) < \mathfrak{n}(\{\mathbf{z}\})$). The only smaller numerosity is 0, so we let $\mathfrak{n}(\{\mathbf{z}\}) = 1$, for any labelled singleton.

(iii) Assume that for a labelled set $\mathbf{A}$, there is a numerosity ξ smaller than the numerosity of the successor $\mathbf{A} \oplus \{\mathbf{y}\}$, so $\mathfrak{n}(\mathbf{A}) < \xi < \mathfrak{n}(\mathbf{A} \oplus \{\mathbf{y}\})$. Then, by the difference principle, there exists $\mathbf{B} \subset \Xi \subset \mathbf{A} \oplus \{\mathbf{y}\}$ with $\mathfrak{n}(\Xi) = \xi$, and

$\mathfrak{n}(\mathbf{B}) = \mathfrak{n}(\mathbf{A})$. Choose $x \in \Xi \setminus B$, and let $\{\mathbf{x}\}$ be the corresponding labelled singleton. Then, by the sum property, and the above, $\mathfrak{n}(\mathbf{B} \oplus \{\mathbf{x}\}) = \mathfrak{n}(\mathbf{A} \oplus \{\mathbf{y}\})$.

But by the definition of $\{\mathbf{x}\}$ and $\oplus$, $|(B \oplus \{x\})_n| \leq |(\Xi)_n|$ for all n, so, by property (i): $\mathfrak{n}(\mathbf{B} \oplus \{\mathbf{x}\}) \leq \mathfrak{n}(\mathbf{\Xi}) < \mathfrak{n}(\mathbf{A} \oplus \{\mathbf{y}\})$, a contradiction.

Let $\mathbf{A} \neq \boldsymbol{\varnothing}$, and $a \in A$ arbitrary. Let $\{\mathbf{a}\}$ be a the labelled singleton with labelling function $\ell_{\{a\}}(a) = \ell_A(a)$. Then for all n, $|(A \setminus \{a\} \oplus \{a\})_n| \leq |A_n|$ and $|A_n| \leq |(A \setminus \{a\} \oplus \{a\})_n|$, so $\mathfrak{n}(\mathbf{A}) = \mathfrak{n}(\mathbf{A} \setminus \{\mathbf{a}\} \oplus \{\mathbf{a}\}) = \mathfrak{n}(\mathbf{A} \setminus \{\mathbf{a}\}) + 1$, so $\mathfrak{n}(\mathbf{A})$ has a predecessor.

(iv) By induction: For $\mathbf{A} = \{\mathbf{a}\}$ a singleton, $\mathfrak{n}(\mathbf{A}) = 1 = |A|$. For $\mathbf{A}$ a labelled set with $|A| = n$, by the above $\mathbf{A}$ has a predecessor $\mathbf{A} \setminus \{\mathbf{a}\}$, with $|A \setminus \{a\}| = n - 1$, so $\mathfrak{n}(\mathbf{A} \setminus \{\mathbf{a}\}) = n - 1$, by hypothesis. Then $\mathfrak{n}(\mathbf{A}) = n - 1 + 1 = n$. □

In the proof of the preceding proposition, the difference principle of the defined numerosity function is heavily used. The consistency of this property will be proven for countable sets. In general, this is yet in question.

3.2 The structure of $\mathscr{N}$

From proposition 3.1.4 it follows that $\mathscr{N}$ can be thought of as having the natural numbers $\mathbb{N}$ as a proper initial segment. The structure of $\mathscr{N}$ is even very much alike.

From property (iii) of definition 3.1.3, it follows that it makes sense to write $\mathfrak{n}(\mathbf{A} \oplus \mathbf{B}) = \mathfrak{n}(\mathbf{A}) + \mathfrak{n}(\mathbf{B})$ and $\mathfrak{n}(\mathbf{A} \odot \mathbf{B}) = \mathfrak{n}(\mathbf{A}) \cdot \mathfrak{n}(\mathbf{B})$. So $\mathscr{N}$ is endowed with a sum and product operation, in the same way as $\mathbb{N}$ is.

The natural numbers $\mathbb{N}$, endowed with the sum and product operations, form a structure that is called a *positive semi-ring with 0 and 1*. It is a semi-ring because it satisfies all axioms of a ring, except for the existence of an additive inverse. It is called *positive* because it is commutative, partially ordered by $\leq$, and it satisfies $x \leq y$ if and only if there is a unique $z \in \mathbb{N}$ such that $x + z = y$.

In the following it will be proven that the set of numerosities $\mathscr{N}$, with the just defined sum and product, has the same structure. This can be done easily by first defining the principle of *isomorphic* labelled sets:

Definition 3.2.1. *Two labelled sets* $\mathbf{A}$ *and* $\mathbf{B}$ *are called* ***isomorphic*** *if there exists a bijective* $\phi : A \to B$ *such that* $\ell_B \circ \phi = \ell_A$. *If this is the case, write* $\mathbf{A} \cong \mathbf{B}$.

From this definition, the following is almost immediate:

Proposition 3.2.2. *(i)* $\mathbf{A} \cong \mathbf{B}$ *iff* $\forall n : |A_n| = |B_n|$

(ii) If $\mathbf{A} \cong \mathbf{B}$, *then* $\mathfrak{n}(\mathbf{A}) = \mathfrak{n}(\mathbf{B})$

(iii) $\mathbf{A} \cong \mathbf{A}'$ *and* $\mathbf{B} \cong \mathbf{B}'$ *gives* $\mathbf{A} \oplus \mathbf{B} \cong \mathbf{A}' \oplus \mathbf{B}'$ *and* $\mathbf{A} \odot \mathbf{B} \cong \mathbf{A}' \odot \mathbf{B}'$

Proof. (i) "⇒": Let $\phi : A \to B$ be an isomorphism. Then $\ell_B \circ \phi = \ell_A$. So $|A_n| = |\{a \in A : \ell_A(a) \leq n\}| = |\{a \in A : \ell_B \circ \phi(a) \leq n\}| = |\{b \in B : \ell_B \circ \phi(\phi^{-1}(b)) \leq n\}| = |\{b \in B : \ell_B(b) \leq n\}| = |B_n|$

"⇐": Assume $|A_n| = |B_n|$. Choose for each $n \in \mathbb{N}$ a bijection $\phi_n : A_n \to B_n$, and let $\phi : A \to B$ be defined by:

$$\phi(a) = \phi_{\ell_A(a)}(a)$$

Then ϕ is an appropriate isomorphism.

(ii) Assume $\mathbf{A} \cong \mathbf{B}$. From the above, it follows $|A_n| \leq |B_n|$ and $|B_n| \leq |A_n|$. Then by definition 3.1.3(i) it follows $\mathfrak{n}(\mathbf{A}) = \mathfrak{n}(\mathbf{B})$.

(iii) Let $\phi : A \to A'$ and $\psi : B \to B'$ be isomorphisms of labelled sets. Then define $\chi : A \oplus B \to A' \oplus B'$ as:

$$\chi(x) = \begin{cases} \phi(x) & \text{if } x \in A \\ \psi(x) & \text{if } x \in B \end{cases}$$

Then χ gives the isomorphism $\mathbf{A} \oplus \mathbf{B} \cong \mathbf{A'} \oplus \mathbf{B'}$, by definition 3.1.2. Define $\eta : A \odot B \to A' \odot B'$ as $\eta(a,b) = (\phi(a), \psi(b))$. Now η gives the isomorphism $\mathbf{A} \odot \mathbf{B} \cong \mathbf{A'} \odot \mathbf{B'}$.

□

With this notion of isomorphic labelled sets, the following proposition is easily proved:

Proposition 3.2.3. *The set of numerosities* $\mathscr{N}$ *endowed with* $+$ *and* $\cdot$ *is a positive semi-ring with* 0 *and* 1.

Proof. In the proof, all ring operations are checked for the labelled set operations $\oplus$ and $\odot$, and then proposition 3.2.2 transfers the isomorphic sets to equal numerosities.

First observe that $\mathbf{A} \oplus \boldsymbol{\varnothing} = \boldsymbol{\varnothing} \oplus \mathbf{A} \cong \mathbf{A}$ (as $\varnothing \uplus A = A$), and for a singleton $\{\mathbf{y}\}$, $\{\mathbf{y}\} \odot \mathbf{A} = \mathbf{A} \odot \{\mathbf{y}\} \cong \mathbf{A}$, with the trivial projection on A as isomorphism. So $\mathfrak{n}(\boldsymbol{\varnothing}) = 0$ and $\mathfrak{n}(\{\mathbf{y}\}) = 1$ serve as 0 and 1.

Commutativity holds: $\mathbf{A} \oplus \mathbf{B} \cong \mathbf{B} \oplus \mathbf{A}$ and $\mathbf{A} \odot \mathbf{B} \cong \mathbf{B} \odot \mathbf{A}$ for obvious reasons.

Associativity follows easily, as $\mathbf{A} \oplus (\mathbf{B} \oplus \mathbf{C}) \cong (\mathbf{A} \oplus \mathbf{B}) \oplus \mathbf{C}$ is trivial from the definition, as is $\mathbf{A} \odot (\mathbf{B} \odot \mathbf{C}) \cong (\mathbf{A} \odot \mathbf{B}) \odot \mathbf{C}$, for $A \times (B \times C) = (A \times B) \times C$ and for $(x,y,z) \in A \times B \times C$:

$$\max\{\max(\ell_A(x), \ell_B(y)), \ell_C(z)\} = \max\{\ell_A(x), \max(\ell_B(y), \ell_C(z))\}$$

Distributivity follows in the same way: $\mathbf{A} \odot (\mathbf{B} \oplus \mathbf{C}) \cong (\mathbf{A} \odot \mathbf{B}) \oplus (\mathbf{A} \odot \mathbf{C})$, and commutativity was already verified.

For $\mathfrak{n}(\mathbf{A}) \leq \mathfrak{n}(\mathbf{B})$, let $\mathbf{A'} \subset \mathbf{B}$ such that $\mathfrak{n}(\mathbf{A'}) = \mathfrak{n}(\mathbf{A})$. Then if $\mathbf{C} = \mathbf{B} \setminus \mathbf{A'}$ (possibly empty) it is obvious that $\mathbf{A'} \oplus \mathbf{C} \cong \mathbf{B}$, thus $\mathfrak{n}(\mathbf{A}) + \mathfrak{n}(\mathbf{C}) = \mathfrak{n}(\mathbf{A} \oplus \mathbf{C}) = \mathfrak{n}(\mathbf{A'} \oplus \mathbf{C}) = \mathfrak{n}(\mathbf{B})$.

Also, for $\mathbf{D}$ a labelled set, $\mathbf{A'} \odot \mathbf{D} \subset \mathbf{B} \odot \mathbf{D}$, and $\mathbf{A'} \oplus \mathbf{D} \subset \mathbf{B} \oplus \mathbf{D}$. So it follows that

$$\mathfrak{n}(\mathbf{A}) + \mathfrak{n}(\mathbf{D}) = \mathfrak{n}(\mathbf{A} \oplus \mathbf{D}) = \mathfrak{n}(\mathbf{A'} \oplus \mathbf{D}) \leq \mathfrak{n}(\mathbf{B} \oplus \mathbf{D}) = \mathfrak{n}(\mathbf{B}) + \mathfrak{n}(\mathbf{D})$$

and $\mathfrak{n}(\mathbf{A}) \cdot \mathfrak{n}(\mathbf{D}) = \mathfrak{n}(\mathbf{A} \odot \mathbf{D}) = \mathfrak{n}(\mathbf{A'} \odot \mathbf{D}) \leq \mathfrak{n}(\mathbf{B} \odot \mathbf{D}) = \mathfrak{n}(\mathbf{B}) \cdot \mathfrak{n}(\mathbf{D})$. So positivity follows. □

So the numerosities $\mathscr{N}$ behave very much like natural numbers, although the set is strictly bigger, and seems to contain “infinite” elements. So what is $\mathscr{N}$?

3.3 A nonstandard universe

Before a sensible answer to the question "what is $\mathscr{N}$?" can be given, a brief introduction of what is called a *nonstandard universe* is needed. The introduction below is based on the introduction to nonstandard analysis in [7] and [3].

From model theory, it is known by the Löwenheim-Skolem theorems that for any consistent theorem in a certain language, there is always a strictly "bigger" model. Applied to the case of Peano arithmetic, for example, which is a set of axioms designed to describe the natural numbers $\mathbb{N}$, this yields a model with a cardinality strictly bigger than that of $\mathbb{N}$ itself.

This is a nice result, but it doesn't tell much about that strictly bigger model, except that one can do Peano arithmetic in the same way as with the natural numbers. It is therefore desirable to do more with these numbers, for example to make statements about *all* functions on these numbers, or even analytical statements, and be able to tell whether these statements hold as well in the "bigger" model.

There is an elegant solution for this, and here is where the language of set theory $L_\in = \{\in\}$ comes in. We'd like to have a model of set theory (ZFC) that satisfies certain statements of first order logic in that language. These statements are in principle statements about sets, but by constructing the model appropriately, the *meaning* of such a statement is a higher order statements about analysis, for example about differentiability or continuity.

Assume this is the goal in the case of one-dimensional real analysis. Then the model should contain $\mathbb{R}$, these are the simplest elements of the model from which everything else is constructed. These elements are called the *urelements*. Suppose one wants to make statements about functions $f : \mathbb{R} \to \mathbb{R}$. Functions are essentially sets (subsets of the cartesian product $\mathbb{R} \times \mathbb{R}$). In the language of set theory an element $(x, y) \in \mathbb{R} \times \mathbb{R}$ is given by the set $\{x, \{x, y\}\} \in \mathscr{P}(\mathscr{P}(\mathbb{R}))$. So for the model to contain *all functions*, it suffices for the model to contain $\mathscr{P}^3(\mathbb{R})$. In order to make statements about differentiability, it has to contain a topology, which is an element of $\mathscr{P}^2(\mathbb{R})$. This way analytical properties can be expressed entirely in the language of set theory, so all that has been proven for first order logic, can be used in the case of these models of set theory.

There is a general procedure for this construction. One wants to have a very large set, called a *superstructure*, constructed from a base set X_0. This set has to be structured in a way such that the language $L_\in = \{\in\}$ of set theory is enough to make statements about higher order mathematics, such as analysis. This superstructure, constructed from a base set X_0, is denoted by $\mathscr{V}(X_0)$. In the example above, the base set is $\mathbb{R}$. The elements of X_0 should certainly be elements of $\mathscr{V}(X_0)$.

Now the definition of the superstructure $\mathscr{V}(X_0)$ has to be made more precise.

Definition 3.3.1 (Superstructure). *For X_0 the base set, define by induction:*

$$X_{n+1} = \mathscr{P}(\bigcup_{k=0}^{n} X_k)$$

Here $\mathscr{P}$ denotes the powerset. Then the superstructure $\mathscr{V}(X_0)$ is defined as:

$$\mathscr{V}(X_0) = \bigcup_{k=0}^{\infty} X_k$$

The set $\mathscr{V}(X_0)$ *is the limit, for* $n \to \infty$ *of the* $n-$*th superstructure approximation* $V_n(X_0)$ *defined as:*

$$V_n(X_0) = \bigcup_{k=0}^{n} X_k$$

Note that every element (*urelement*) of the base set X_0 is contained in $\mathscr{V}(X_0)$. The set elements of $\mathscr{V}(X_0)$ are called *entities*, adopting the terminology from [7].

There are a few properties of $\mathscr{V}(X_0)$ that are worth mentioning, as they give this superstructure enough sets to talk about higher order mathematics, such as continuous functions, differentiability, etcetera. They are stated in the following proposition. Like above, with (x_1, x_2) is meant the *ordered tuple* $\{\{x\},\{x,y\}\}$, and by induction $(x_1,\dots,x_n) = ((x_1,\dots,x_{n-1}),x_n)$. The cartesian product of n sets is defined as the set of all such n-tuples.

Proposition 3.3.2. *For* $\mathscr{V}(X_0)$ *a superstructure with base set* X_0*, any* $x_1,\dots x_n \in \mathscr{V}(X_0)$*, and entities* $u, v \in \mathscr{V}(X_0)$*:*

(i) $\{x_1,\dots,x_n\} \in \mathscr{V}(X_0)$

(ii) $(x_1,\dots,x_n) \in \mathscr{V}(X_0)$

(iii) For all $w \subset u$, $w \in \mathscr{V}(X_0)$

(iv) $u \times v \in \mathscr{V}(X_0)$

(v) $\mathscr{P}(u) \in \mathscr{V}(X_0)$

Proof. (i) For some k, $x_i \in X_k$ for all $1 \le i \le n$. Then $\{x_1,\dots x_n\} \in \mathscr{P}(X_k) \subset X_{k+1} \subset \mathscr{V}(X_0)$. (ii) For any $x, y \in \mathscr{V}(X_0)$, $x, y \in X_k$ for some k. Then $(x,y) = \{\{x\},\{x,y\}\} \in \mathscr{P}(\mathscr{P}(X_k)) \subset X_{k+2} \subset \mathscr{V}(X_0)$. By induction, this holds for all n-tuples. (iii) $u \in X_k$ for some k. Then $u \subset X_{k-1}$ so $w \subset X_{k-1}$. Then $w \in \mathscr{P}(X_{k-1}) \subset X_k \subset \mathscr{V}(X_0)$. (iv) For some $k \in \mathbb{N}$, $u, v \in X_k$. $u \times v = \{(x,y) | x \in u, y \in v\} \subset \{(x,y) | x, y \in X_{k-1}\} \subset \mathscr{P}(X_{k+1})$ by (ii). So then $u \times v \in \mathscr{P}(\mathscr{P}(X_{k+1})) \subset X_{k+3} \subset \mathscr{V}(X_0)$. (v) $u \in X_k$, so $u \subset X_{k-1}$, so for all $w \subset u$, $w \subset X_{k-1}$, hence $w \in X_k$. Then it follows $\mathscr{P}(u) = \{w \subset u\} \subset \mathscr{P}(X_k) \subset X_{k+1} \subset \mathscr{V}(X_0)$. □

Remark that $\mathscr{V}(X_0)$ is transitive set. From the above, it follows that $\mathscr{V}(X_0)$ is closed under basic analytical and topological constructions on the base set. For example, if $X_0 = \mathbb{R}$ is used as a based set, it is possible that one could make statements such as:

$$\phi(f,a) = (\forall \epsilon > 0)\exists \delta > 0 : \forall (x_1,\dots,x_n) \in B_\delta(a), f(x_1,\dots,x_n) \in B_\epsilon(f(a))$$

This statement expresses continuity of f in the point a in the language of set theory. Note that all used sets, such as functions, open balls in $\mathbb{R}^n$, are all contained in the superstructure $\mathbb{R}$. So $\phi(f,a)$ is expressible entirely in the language of set theory.

The goal is now to be able to have another superstructure $\mathscr{V}(Y_0)$ which is bigger than $\mathscr{V}(X_0)$, and a *bounded elementary embedding* $* : \mathscr{V}(X_0) \to \mathscr{V}(Y_0)$. A *bounded elementary embedding* is an embedding in the usual sense, so it induces an isomorphism $\mathscr{V}(X_0) \cong \mathrm{Im}_*(\mathscr{V}(X_0))$ between its domain and image. It differs from an elementary embedding in the fact that it transfers only *bounded*

$L_\in$-sentences. A *bounded* $L_\in$ sentence is a sentence ϕ whose quantifiers always quantify over a set. These quantifiers are called *bounded quantifiers* and are always of the form $(\forall x \in y)\ldots$ or $(\exists x \in y)\ldots$, which is short for $(\forall x) x \in y \rightarrow \ldots$ or $(\exists x) x \in y \wedge \ldots$ respectively.

For instance, $\phi(f, a)$ is in fact a bounded sentence, because with $\forall \epsilon > 0$ is meant $\forall \epsilon \in R_{>0}$, and in the same way every quantifier is bounded. This property must therefore, under a bounded elementary embedding $*$, also hold in $\mathscr{V}(Y_0)$, that is:

$$\mathscr{V}(X_0) \vDash \phi(f, a) \iff \mathscr{V}(Y_0) \vDash \phi({}^*f, {}^*a)$$

In short, $*$ satisfies the *Leibniz' transfer principle*:

Definition 3.3.3 (Leibniz' transfer principle)**.** *For $\mathscr{V}(X_0)$ and $\mathscr{V}(Y_0)$ $L_\in$-models, a map $* : \mathscr{V}(X_0) \rightarrow \mathscr{V}(Y_0)$ satisfies the Leibniz transfer principle if it is an $L_\in$-embedding and for every bounded $L_\in$-formula $\psi[\xi_1, \ldots, \xi_n]$, and $x_1, \ldots, x_n \in \mathscr{V}(X_0)$ holds: $\mathscr{V}(X_0) \vDash \psi[x_1, \ldots, x_n] \iff \mathscr{V}(Y_0) \vDash \psi[*x_1, \ldots, *x_n]$. Such a $*$ is called a* ***bounded elementary embedding****.*

To avoid ambiguity in the following definition, the image of a set $A \subseteq \mathscr{V}(X_0)$ under $*$ will always be denoted as $\mathrm{Im}_*(A)$, like is done in the case above. The image of an element $A \in \mathscr{V}(X_0)$ (which might be an internal set) under $*$ is denoted as ${}^*A$. In this notation, $\mathrm{Im}_*(A) = \{{}^*a : a \in A\} \neq {}^*A$ for arbitrary $A \in \mathscr{V}(X_0)$.

Now the following defines a *nonstandard universe*:

Definition 3.3.4. *By a* ***nonstandard universe*** *is meant a triple $\langle \mathscr{V}(X_0), \mathscr{V}(Y_0), * \rangle$ such that:*

(i) $* : \mathscr{V}(X_0) \rightarrow \mathscr{V}(Y_0)$ satisfies the Leibniz' transfer principle, i.e. it is a bounded elementary embedding.

(ii) ${}^*X_0 = Y_0$

(iii) For every infinite subset $U \subset X_0$, the set $Im_*(U)$ is a proper subset of ${}^*U$. A set $Y \in \mathscr{V}(Y_0)$ is called ***internal*** *if there is a set $X \in \mathscr{V}(X_0)$ such that $Y \in {}^*X$.*

If these three properties are satisfied, $*$ is called a non-standard embedding.

It should be clear that this result is only useful if many properties are expressible using bounded $L_\in$-sentences. That is the case, examples can be found in [3].

The following definition gives rise to a certain type of non-standard embeddings:

Definition 3.3.5. *A family of sets $\mathscr{Y}$ is said to have the* ***finite intersection property*** *if for every finite subfamily $\mathscr{Y}' \subset \mathscr{Y}$ holds $\cap \mathscr{Y}' \neq \varnothing$.*

Definition 3.3.6. *Let $\langle \mathscr{V}(X_0), \mathscr{V}(Y_0), * \rangle$ be a non-standard universe. If for every countable family $\mathscr{Y}$ of internal sets with finite intersection property holds that $\cap \mathscr{Y} \neq \varnothing$, $*$ is called $\omega-$**saturated**.*

The question that now arises is: Under what conditions do nonstandard universes exist? In the following section, it will be proven that existence of a numerosity function provides a nonstandard universe $\langle \mathscr{V}(\mathbb{N}), \mathscr{V}(\mathscr{N}), * \rangle$, where $*$ is an ω-saturated nonstandard embedding.

3.4 From numerosities to hyperintegers

The approximating sequences $n \mapsto |A_n|$ for labelled sets $\mathbf{A}$ can be identified with the set $\mathscr{F}$ of nondecreasing sequences in $\mathbb{N}$:

$$\mathscr{F} = \{\phi \in \mathbb{N}^{\mathbb{N}} | i \le j \implies \phi_i \le \phi_j\}$$

Every approximating sequence $(|A_n|)_{n \in \mathbb{N}}$ is indeed a nondecreasing sequence in $\mathbb{N}$. Furthermore, for $\phi \in \mathscr{F}$, define by induction:

$$\begin{aligned} A_{\phi,0} &= \{0, 1, \dots, \phi_0 - 1\} \\ A_{\phi,n} &= \{0, 1, \dots, \phi_n - \phi_{n-1} - 1\} \quad \text{for } n \ge 1 \end{aligned}$$

Where if $\phi_n - \phi_{n-1} - 1 < 0$, $A_{\phi,n} = \varnothing$. Then let $\mathbf{A}_\phi$ be the labelled set with $A_\phi = \uplus_{n=0}^{\infty} A_{\phi,n}$, and $\ell_{A_\phi}(a) = i \iff a \in A_{\phi,i}$. Observe ϕ is the approximating sequence of $\mathbf{A}_\phi$:

$$|(A_\phi)_n| = |\{a \in A_\phi | \ell_{A_\phi}(a) \le n\}| = |\cup_{k=0}^{n} A_{\phi,k}| = \phi_0 + \sum_{k=1}^{n} (\phi_k - \phi_{k-1}) = \phi_n$$

Remark the difference in notation: $A_{\phi,n}$ is the set defined above, $(A_\phi)_n$ is the n−th finite approximation of $\mathbf{A}_\phi$.

From the above it follows, by inclusion in both directions, $\mathscr{F}$ is the set of all approximating sequences of labelled sets. With the canonical sum and product operations $(\psi \cdot \phi)_n = \psi_n \cdot \phi_n$, $(\psi + \phi)_n = \psi_n + \phi_n$, and ordering $\psi \le \phi \iff \forall n : \psi_n \le \phi_n$, $\langle \mathscr{F}, +, \cdot, \le \rangle$ is a partially ordered semi-ring, as is $\mathscr{N}$. It is not positive though:

$$\begin{aligned} &\text{for } \phi_n = 1, \psi_0 = 0 \text{ and } \psi_n = 1 \forall n \ge 1 : \\ &\psi \le \phi \text{ but there is no } \chi \in \mathscr{F} \text{ such that } \psi + \chi = \phi \end{aligned}$$

The following lemma, however, proves a homomorphism between the two semi-rings.

Lemma 3.4.1. *The map $\rho : \mathscr{F} \twoheadrightarrow \mathscr{N}$ given by $\rho(\phi) = \mathfrak{n}(\mathbf{A}_\phi)$, with A_ϕ defined by the construction above, is a homomorphism of partially ordered semi-rings.*

Proof. Define a bijection $f : A_\phi \uplus A_\psi \to A_{\phi+\psi}$ by choosing bijections $f_n : A_{\phi,n} \uplus A_{\psi,n} \to A_{\phi+\psi,n}$ and defining $f(a) = f_n(a) \iff a \in A_{\phi,n} \uplus A_{\psi,n}$. Each bijection exists, as $|A_{\phi,n} \uplus A_{\psi,n}| = |A_{\phi+\psi,n}| = \phi_n + \psi_n - \phi_{n-1} - \psi_{n-1}$. From the definition it immediately follows that f induces an isomorphism of labelled sets $\mathbf{A}_\phi \oplus \mathbf{A}_\psi \cong \mathbf{A}_{\phi+\psi}$. So $\rho(\phi + \psi) = \mathfrak{n}(\mathbf{A}_{\phi+\psi}) = \mathfrak{n}(\mathbf{A}_\phi \oplus \mathbf{A}_\psi) = \mathfrak{n}(\mathbf{A}_\phi) + \mathfrak{n}(\mathbf{A}_\psi) = \rho(\phi) + \rho(\psi)$.

In the same way, because $A_\phi \times A_\psi = \uplus_{n=0}^{\infty} A_{\phi,n} \times \uplus_{n=0}^{\infty} A_{\psi,n} = \uplus_{n=0}^{\infty} A_{\phi,n} \times A_{\psi,n}$ (just rearrange the elements),the set of $g_n : A_{\phi,n} \times A_{\psi,n} \to A_{\phi \cdot \psi,n}$ provides the bijections to produce an isomorphism $\mathbf{A}_\phi \odot \mathbf{A}_\psi \cong \mathbf{A}_{\phi \cdot \psi}$ which gives $\rho(\phi \cdot \psi) = \mathfrak{n}(\mathbf{A}_{\phi \cdot \psi}) = \mathfrak{n}(\mathbf{A}_\phi \odot \mathbf{A}_\psi) = \mathfrak{n}(\mathbf{A}_\phi) \cdot \mathfrak{n}(\mathbf{A}_\psi) = \rho(\phi) \cdot \rho(\psi)$.

Finally, $\phi \le \psi \implies \forall n : |A_{\phi,n}| \le |A_{\psi,n}| \implies \forall n : |(A_\phi)_n| \le |(A_\psi)_n| \implies \rho(\phi) = \mathfrak{n}(\mathbf{A}_\phi) \le \mathfrak{n}(\mathbf{A}_\psi) = \rho(\psi)$ by definition 3.1.3(i). $\square$

The construction of the intended embedding $* : \mathscr{V}(\mathbb{N}) \to \mathscr{V}(\mathscr{N})$ is very much like the common ultrapower construction found in literature. At this point, it is namely possible to define a set $\mathfrak{U} \subset \mathscr{P}(\mathbb{N})$ using ρ, which will be proven to be an ultrafilter on $\mathbb{N}$.

Proposition 3.4.2. *Let* $\mathfrak{U} = \{\{n : \phi_n = \psi_n\} | \psi, \phi \in \mathscr{F}$ *with* $\rho(\phi) = \rho(\psi)\}$*. Then the following holds:*

(i) $D = \{n : \phi_n = \psi_n\} \in \mathfrak{U} \iff \rho(\phi) = \rho(\psi)$.

(ii) $D, E \in \mathfrak{U}$ *implies* $D \cap E \in \mathfrak{U}$.

(iii) $D \in \mathfrak{U} \iff D^c \notin \mathfrak{U}$.

(iv) If $D \in \mathfrak{U}$ *and* $E \supset D$*, then* $E \in \mathfrak{U}$.

(v) If D *is finite, then* $D \notin \mathfrak{U}$.

In [1], elements of $\mathfrak{U}$ are called *qualified* sets. The proposition can be proven with the following, very useful lemma:

Lemma 3.4.3. *(i)* $\varnothing \notin \mathfrak{U}$

(ii) Let $\vartheta_D \in \mathscr{F}$ *be the function such that:*

$$\vartheta_D(n) = \begin{cases} n & \textit{if } n \in D \\ n+1 & \textit{otherwise} \end{cases}$$

Then $D \in \mathfrak{U}$ *iff* $\rho(\vartheta_D) = \rho(1_{\mathbb{N}}) = \alpha$*, where* $1_{\mathbb{N}}$ *is the identity on* $\mathbb{N} : n \mapsto n$

Proof. (i) We have to prove that if $\forall n : \phi_n \neq \psi_n$, then $\rho(\phi) \neq \rho(\psi)$. For such ϕ, ψ, we have for all n:

$$\begin{array}{lll} (\phi_n - \psi_n)^2 & > & 0 \\ \phi_n^2 + \psi_n^2 & > & 2\phi_n\psi_n \\ \phi_n^2 + \psi_n^2 & \geq & 2\phi_n\psi_n + 1 \end{array}$$

From the last line follows $\rho(\phi)^2 + \rho(\psi)^2 \geq 2\rho(\phi)\rho(\psi) + 1$, so $\rho(\phi) \neq \rho(\psi)$.

(ii) It is obvious that if $\rho(\vartheta_D) = \rho(1_{\mathbb{N}})$, then by definition $\{n : \vartheta_D(n) = 1_{\mathbb{N}}(n)\} = D \in \mathfrak{U}$. For the converse, assume that $D \in \mathfrak{U}$ is of the form $D = \{n : \tilde{\phi}_n = \tilde{\psi}_n\}$ for $\tilde{\phi}, \tilde{\psi}$ nondecreasing with $\rho(\tilde{\phi}) = \rho(\tilde{\psi})$.

Define $\phi = \tilde{\phi} + 1_{\mathbb{N}}, \psi = \tilde{\psi} + 1_{\mathbb{N}}$. Then $D = \{n : \phi_n = \psi_n\}$ with ϕ, ψ strictly increasing and $\rho(\phi) = \rho(\psi)$.

Now let $\tau = \phi + \vartheta_D - 1_{\mathbb{N}}$. Then $\tau_n = \phi_n$ if $n \in D$, $\tau_n = \phi_n + 1$ otherwise. Then for all n we have $\tau_n \neq \psi_n + 1$, so $\rho(\tau) \neq \rho(\psi + 1) = \rho(\psi) + 1$, by (i). We now have:

$$\begin{array}{l} \phi \leq \tau \leq \phi + 1 \\ \rho(\phi) \leq \rho(\tau) \leq \rho(\phi) + 1 = \rho(\psi) + 1, \text{but } \rho(\tau) \neq \rho(\psi) + 1 \text{ so:} \\ \rho(\psi) = \rho(\phi) \leq \rho(\tau) < \rho(\psi) + 1 \\ \rho(\phi) = \rho(\tau) = \rho(\phi) + \rho(\vartheta_D) - \rho(1_{\mathbb{N}}) \\ \rho(\vartheta_D) = \rho(1_{\mathbb{N}}) \end{array}$$

□

The proof of proposition 3.4.2 follows:

Proof. (i) "⇐" is by definition. For "⇒", assume $D = \{n : \phi_n = \psi_n\} \in \mathfrak{U}$. With the notation of lemma 3.4.3, observe:

$$\begin{aligned}
(\phi \cdot 1_{\mathbb{N}} + \phi + \psi \cdot \vartheta_D)_n &= \begin{cases} (n+1)\phi_n + n\psi_n = (n+1)\psi_n + n\phi_n \text{ if } n \in D \\ (n+1)\phi_n + (n+1)\psi_n \text{ if } n \notin D \end{cases} \\
&= (\psi \cdot 1_{\mathbb{N}} + \psi + \phi \cdot \vartheta_D)_n \\
\phi \cdot 1_{\mathbb{N}} + \phi + \psi \cdot \vartheta_D &= \psi \cdot 1_{\mathbb{N}} + \psi + \phi \cdot \vartheta_D
\end{aligned}$$

Applying ρ to the above, using lemma 3.4.3(ii) then yields:

$$\begin{aligned}
\rho(\phi) \cdot \rho(1_{\mathbb{N}}) + \rho(\phi) + \rho(\psi) \cdot \rho(\vartheta_D) &= \rho(\psi) \cdot \rho(1_{\mathbb{N}}) + \rho(\psi) + \rho(\phi) \cdot \rho(\vartheta_D) \\
\rho(\phi + \psi) \cdot \alpha + \rho(\phi) &= \rho(\phi + \psi) \cdot \alpha + \rho(\psi) \\
\rho(\phi) &= \rho(\psi)
\end{aligned}$$

Where $\alpha = \rho(1_{\mathbb{N}})$.

(ii) Note that, if $D, E \in \mathfrak{U}$:

$$\rho(\vartheta_D \cdot \vartheta_E) = \rho(\vartheta_D) \cdot \rho(\vartheta_E) = \rho(1_{\mathbb{N}})^2 = \rho(n \mapsto n^2)$$

So this implies $\{n : \vartheta_D(n) \cdot \vartheta_E(n) = n^2\} \in \mathfrak{U}$, but this set is precisely the set of n such that $\vartheta_D(n) = \vartheta_E(n) = n$, i.e. the set $D \cap E$.

(iii) "⇒" follows immediately, as if $D^c \in \mathfrak{U}$, then from property (ii) above we have $D^c \cap D = \varnothing \in \mathfrak{U}$ which contradicts lemma 3.4.3(i).

For "⇐", note $\vartheta_D + \vartheta_{D^c} = 1_{\mathbb{N}} + 1_{\mathbb{N}} + 1$. Remark that $1_{\mathbb{N}} \leq \vartheta_{D^c} \leq 1_{\mathbb{N}} + 1$. By Assumption, and lemma 3.4.3, $\rho(\vartheta_{D^c}) \neq \rho(1_{\mathbb{N}})$. So, ρ being a homomorphism of partially ordered semi-rings, it follows:

$$\rho(1_{\mathbb{N}}) < \rho(\vartheta_{D^c}) \leq \rho(1_{\mathbb{N}}) + 1$$

So $\rho(\vartheta_{D^c}) = \rho(1_{\mathbb{N}}) + 1$. Therefore it must hold that $\rho(\vartheta_D) = \rho(1_{\mathbb{N}})$, hence $D \in \mathfrak{U}$.

(iv) If $E \notin \mathfrak{U}$, then by the above $E^c \in \mathfrak{U}$, hence $\mathfrak{U} \ni D \cap E^c = \varnothing$, this contradicts lemma 3.4.3(i).

(v) Let $D \subset \mathbb{N}$ be finite, i.e. $D = \{i_1, \ldots, i_k\}$ with $i_1 < \cdots < i_k$, and assume $D \in \mathfrak{U}$. The goal is to find approximating sequences ϕ, ψ such that $D = \{n : \phi_n = \psi_n\}$, and arrive at a contradiction. Define the labelled sets $\mathbf{A}, \mathbf{B}$ as follows:

$$\begin{aligned}
A &= \{0, 1, \ldots, k\} \quad \ell_A(0) = 0, \ell_A(j) = i_j + 1 \text{ for } 1 \leq j \leq k \\
B &= \{1, 2, \ldots, k\} \quad \ell_B(j) = i_j
\end{aligned}$$

And let $\phi_n = |A_n|$,$\psi_n = |B_n|$ be the approximating sequences. Observe for all j:

$$\begin{aligned}
|A_{i_j}| &= 1 + (j - 1) = j \\
|B_{i_j}| &= j
\end{aligned}$$

And for all $i_j < l < i_{j+1}$:

$$\begin{aligned}
|A_l| &= j + 1 \\
|B_l| &= j
\end{aligned}$$

While for $l < i_1 : |A_l| = 1, |B_l| = 0$, and for $l > i_k : |A_l| = k+1, |B_l| = k$. It follows that $D = \{n : \phi_n = \psi_n\}$, but $\rho(\phi) = \mathfrak{n}(\mathbf{A}) = k+1 \neq k = \mathfrak{n}(\mathbf{B}) = \rho(\psi)$. This contradicts property (i).

□

The *ultrafilter* properties of $\mathfrak{U}$ follow immediately from the previous proposition and lemma. So, in the notation of appendix A.1.1, what is proven so far is that the set of numerosities $\mathscr{N}$ is contained in an ultraproduct of $\mathbb{N}$:

$$\mathscr{N} = \{\rho(\phi) | \phi \in \mathscr{F}\} = \{\phi \in \mathscr{F}\}/\mathfrak{U} \subseteq \prod_{\mathfrak{U}} \mathbb{N}$$

The following proposition even leads to equality of the above:

Proposition 3.4.4. *For every* $\phi : \mathbb{N} \to \mathbb{N}$*, there is a* $D \in \mathfrak{U}$ *such that* $\phi|_D$ *is nondecreasing.*

Proof. We have to prove that for any such ϕ there must be a $\psi \in \mathscr{F}$ such that $D = \{n : \psi_n = \phi_n\} \in \mathfrak{U}$. Let now $\phi : \mathbb{N} \to \mathbb{N}$ be arbitrary.

Define $\tau, \chi \in \mathscr{F}$ by $\tau(n+1) = \chi(n) = \sum_{i \le n} \phi_n$. First remark that $\phi = \chi - \tau$. Notice $\tau \le \chi$, so $\rho(\tau) \le \rho(\chi)$, hence there exists $\mathbf{A} \subseteq \mathbf{B}$ such that:

$$\rho(\tau) = \mathfrak{n}(\mathbf{A}) \le \mathfrak{n}(\mathbf{B}) = \rho(\chi)$$

Let $\mathbf{C} = \mathbf{B} \setminus \mathbf{A}$ the labelled set (with the restricted labelling from $\mathbf{B}$), and let $\psi_n = |C_n|$ be the approximating sequence. Then:

$$\rho(\psi + \tau) = \rho(\psi) + \rho(\tau) = \mathfrak{n}(\mathbf{C}) + \mathfrak{n}(\mathbf{A}) = \mathfrak{n}(\mathbf{B}) = \rho(\chi)$$

So by property (i) of proposition 3.4.2 it follows $D = \{n : \psi_n + \tau_n = \chi_n\} = \{n : \psi_n = \phi_n\} \in \mathfrak{U}$, which was to be proven. □

This leads to the following corollary:

Corollary 3.4.5. *With* $\mathfrak{U}$ *as in proposition 3.4.2, we have:*

$$\mathscr{N} = \prod_{\mathfrak{U}} \mathbb{N}$$

This actually gives an insightful view of what an ultrapower of a "standard" base set looks like, in a semantic way. In this case, the ultrapower construction extends the natural numbers to a set of numerosities, that have a "meaning" themselves (that is, sizes of labelled sets).

The propositions above are of great importance for the next proposition. By themselves, they again show the very natural behaviour of numerosities. The above ultraproduct can already be seen as a bigger (non-standard) model $\mathbb{N}$, as is argued in the appendix. But there is more, the goal was to transfer bounded $L_\in$-sentences from the superstructure $\mathscr{V}(\mathbb{N})$ to $\mathscr{V}(\mathscr{N})$. That will be done using one more proposition, but first I introduce some notation.

Using definition 3.3.1, define $\mathscr{F}_n = \{\phi | \phi : \mathbb{N} \to V_n(\mathbb{N})\}$. These sets form an increasing chain:

$$\mathscr{F}_0 \subset \mathscr{F}_1 \subset \mathscr{F}_2 \subset \ldots$$

Define $\mathscr{F}_\infty = \cup_{n=0}^{\infty} \mathscr{F}_n = \{\phi : \phi : \mathbb{N} \to V_k(\mathbb{N}) \text{ for some } k\}$.

With the tools above, the goal is to extend $\rho : \mathscr{F} \to \mathscr{N}$ to a map $\boldsymbol{\rho} : \mathscr{F}_\infty \to \mathscr{V}(\mathscr{N})$. This map can be easily used to define the nonstandard embedding $* : \mathscr{V}(\mathbb{N}) \to \mathscr{V}(\mathscr{N})$.

To improve readability, we introduce the following relations for $\phi, \psi \in \mathscr{F}_\infty$: $\phi =_{\mathfrak{U}} \psi$ if $\{n : \phi_n = \psi_n\} \in \mathfrak{U}$, $\phi \in_{\mathfrak{U}} \psi$ if $\{n : \phi_n \in \psi_n\} \in \mathfrak{U}$. Note that "$=_{\mathfrak{U}}$" is an equivalence relation. Denote by ${}^cA \in \mathscr{F}_\infty$ the constant sequence ${}^cA(n) = A$, for A a set. The following proposition gives the definition of $\boldsymbol{\rho}$:

Proposition 3.4.6. *There exists a unique map* $\boldsymbol{\rho} : \mathscr{F}_\infty \to \mathscr{V}(\mathscr{N})$ *such that:*

(i) For all $\phi \in \mathscr{F}$, $\boldsymbol{\rho}(\phi) = \rho(\phi)$

(ii) $\boldsymbol{\rho}({}^c\varnothing) = \varnothing$

(iii) $\boldsymbol{\rho}(\phi) = \boldsymbol{\rho}(\psi) \iff \phi =_{\mathfrak{U}} \psi$

(iv) $\boldsymbol{\rho}(\phi) \in \boldsymbol{\rho}(\psi) \iff \phi \in_{\mathfrak{U}} \psi$

Proof. For $\phi =_{\mathfrak{U}} {}^c\varnothing$, let $\boldsymbol{\rho}(\phi) = \varnothing$.

For $\phi \in \mathscr{F}_\infty$, $\phi \neq_{\mathfrak{U}} {}^c\varnothing$, the proof is by induction on $k = \min\{h : \phi \in_{\mathfrak{U}} {}^cV_h(\mathbb{N})\}$. So first assume $k = 0$, that is $\phi \in {}^c\mathbb{N}$. Then, by proposition 3.4.4, there exists $\phi' \in \mathscr{F}$ such that $\phi =_{\mathfrak{U}} \phi'$. Now define $\boldsymbol{\rho}(\phi) = \rho(\phi')$. This way properties (i)-(iii) follow immediately for $k = 0$.

For $k > 0$, note that $\chi \in_{\mathfrak{U}} \phi$ implies $\chi \in_{\mathfrak{U}} {}^cV_{k-1}(\mathbb{N})$. Then define $\boldsymbol{\rho}(\phi) = \{\boldsymbol{\rho}(\chi) | \chi \in_{\mathfrak{U}} \psi\}$ inductively. Now only properties (iii) and (iv) and uniqueness are to be proven.

For (iii), let $l = \max\{\min\{h : \phi \in_{\mathfrak{U}} {}^cV_h(\mathbb{N})\}, \min\{h : \psi \in_{\mathfrak{U}} {}^cV_h(\mathbb{N})\}\}$, and observe (iii) already holds for $l = 0$. Now assume it holds for all $i < l$.

For one direction, assume $\phi =_{\mathfrak{U}} \psi$. Then, for $\chi \in_{\mathfrak{U}} \phi$ arbitrary, it follows from proposition 3.4.2 (ii) and (iv) that:

$$\begin{aligned} &\{n : \chi_n \in \psi_n\} \supset \{n : \phi_n = \psi_n\} \cap \{n : \chi_n \in \phi_n\} \in \mathfrak{U} \text{ so} \\ &\{n : \chi_n \in \psi_n\} \in \mathfrak{U} \end{aligned}$$

So $\chi \in_{\mathfrak{U}} \psi$. The converse also holds, so $\chi \in_{\mathfrak{U}} \phi \iff \chi \in_{\mathfrak{U}} \psi$. By definition, it now follows $\boldsymbol{\rho}(\phi) = \boldsymbol{\rho}(\psi)$. By induction, this holds in general.

For the converse, assume $\phi \neq_{\mathfrak{U}} \psi$. Then by proposition 3.4.2 (iii) it follows $\{n : \phi_n \neq \psi_n\} \in \mathfrak{U}$. Then there exist $\chi, \xi \in_{\mathfrak{U}} {}^cV_{l-1}$ such that $\chi \in_{\mathfrak{U}} \phi$, $\xi \in_{\mathfrak{U}} \psi$, but $\chi \neq_{\mathfrak{U}} \xi$: For example, let χ, ξ be such that for all $n \in \{n : \phi_n \neq \psi_n\} \in \mathfrak{U}$:

$$\chi_n \in \phi_n \setminus \psi_n \text{ or } \xi_n \in \psi_n \setminus \phi_n \text{ and } \chi_n \in \phi_n, \xi_n \in \psi_n \text{ in all cases}$$

Then obviously $\chi \in_{\mathfrak{U}} \phi$ and $\xi \in_{\mathfrak{U}} \psi$, but $\{n : \chi_n = \xi_n\} \subseteq \{n : \phi_n = \psi_n\} \notin \mathfrak{U}$, so $\chi \neq_{\mathfrak{U}} \xi$. By hypothesis (it holds for all $i < l$), $\boldsymbol{\rho}(\chi) \neq \boldsymbol{\rho}(\xi)$, so by definition it now follows $\boldsymbol{\rho}(\phi) \neq \boldsymbol{\rho}(\psi)$.

For (iv), "$\Leftarrow$" is by definition. For "$\Rightarrow$", assume $\boldsymbol{\rho}(\phi) \in \boldsymbol{\rho}(\psi)$. Then there exists $\phi' \in_{\mathfrak{U}} \psi$ such that $\boldsymbol{\rho}(\phi) = \boldsymbol{\rho}(\phi')$. By (iii), $\phi =_{\mathfrak{U}} \phi'$, so then $\phi \in_{\mathfrak{U}} \psi$. By induction, this holds in general.

Uniqueness follows from property (i) and property (iv), in other words $\boldsymbol{\rho}$ is entirely defined by ρ. ☐

In the following theorem, the map $* : x \mapsto \boldsymbol{\rho}({}^cx)$ will be proven to be a nonstandard embedding $\mathscr{V}(\mathbb{N}) \to \mathscr{V}(\mathscr{N})$. To make things clear, it is important to realize that the construction above is basically the ultrapower construction as found in [3]. There, using an ultrafilter D on a index set I, and X a base set, a nonstandard base set Y is constructed as the ultrapower: $Y = \prod_D X$ (see appendix A.1.1). The *natural embedding* d of X into Y is given by $d : x \mapsto [{}^c x]_D$ where ${}^c x(i) = x$ is the constant sequence with value x for all $i \in I$.

The embedding $* : x \mapsto \boldsymbol{\rho}({}^c x) = [{}^c x]_{\mathfrak{U}}$ is analogous, by corollary 3.4.5. Due to the analogy in construction, the proof of the following theorem is very much like the proof of existence of a nonstandard embedding by the ultrapower construction.

Theorem 3.4.7. *The map $* : \mathscr{V}(\mathbb{N}) \to \mathscr{V}(\mathscr{N})$ given by $* : x \mapsto \boldsymbol{\rho}({}^c x)$ is an ω-saturated nonstandard embedding (recall definition 3.3.6). The set of internal elements is precisely the range of $\boldsymbol{\rho}$.*

Proof. First we prove that $*$ is a bounded elementary embedding. It must be proven that for $\sigma[\xi_1, \dots, \xi_n]$ a bounded $L_\in$-formula, $x_1, \dots, x_n \in \mathscr{V}(\mathbb{N})$:

$$\mathscr{V}(\mathbb{N}) \models \sigma[x_1, \dots, x_n] \iff \mathscr{V}(\mathscr{N}) \models \sigma[*x_1, \dots, *x_n]$$

This is done by first proving the following intermediate result, where $\phi^1, \dots, \phi^n \in \mathscr{F}_\infty$:

$$(\star)\ \ \{i : \mathscr{V}(\mathbb{N}) \models \sigma[\phi^1(i), \dots, \phi^n(i)]\} \in \mathfrak{U} \iff \mathscr{V}(\mathscr{N}) \models \sigma[\boldsymbol{\rho}(\phi^1), \dots, \boldsymbol{\rho}(\phi^n)]$$

Proof ($\star$). The proof is by induction on the complexity of $L_\in$-formulas. It holds for atomic formulas:

$$\begin{aligned} \phi^i =_{\mathfrak{U}} \phi^j &\iff \boldsymbol{\rho}(\phi^i) = \boldsymbol{\rho}(\phi^j) \\ \phi^i \in_{\mathfrak{U}} \phi^j &\iff \boldsymbol{\rho}(\phi^i) \in \boldsymbol{\rho}(\phi^j) \end{aligned}$$

As follows directly from proposition 3.4.6 (iii) and (iv). Now there's still to prove that the set of $L_\in$-formulas for which $(\star)$ holds is closed under concatenation by $\wedge$, negation $\neg$, quantifying over an element $\exists x \in y$, and equivalence of formulas. Suppose $(\star)$ holds for an $L_\in$ formula σ. Then:

$$\begin{aligned} &\{i : \mathscr{V}(\mathbb{N}) \models \neg\sigma[\phi^1(i), \dots, \phi^n(i)]\} = \{i : \mathscr{V}(\mathbb{N}) \models \sigma[\phi^1(i), \dots, \phi^n(i)]\}^c \in \mathfrak{U} \iff \\ &\{i : \mathscr{V}(\mathbb{N}) \models \sigma[\phi^1(i), \dots, \phi^n(i)]\} \notin \mathfrak{U} \iff \mathscr{V}(\mathscr{N}) \not\models \sigma[\boldsymbol{\rho}(\phi^1), \dots, \boldsymbol{\rho}(\phi^n)] \iff \\ &\mathscr{V}(\mathscr{N}) \models \neg\sigma[\boldsymbol{\rho}(\phi^1), \dots, \boldsymbol{\rho}(\phi^n)] \end{aligned}$$

By the properties of $\mathfrak{U}$.

Suppose $(\star)$ holds for $L_\in$ formulas σ_i, σ_j. Then:

$$\begin{aligned} &\{i : \mathscr{V}(\mathbb{N}) \models \sigma_i \wedge \sigma_j[\phi^1(i), \dots, \phi^n(i)]\} = \\ &\{i : \mathscr{V}(\mathbb{N}) \models \sigma_i[\phi^1(i), \dots, \phi^n(i)]\} \cap \{i : \mathscr{V}(\mathbb{N}) \models \sigma_j[\phi^1(i), \dots, \phi^n(i)]\} \in \mathfrak{U} \iff \\ &\{i : \mathscr{V}(\mathbb{N}) \models \sigma_i[\phi^1(i), \dots, \phi^n(i)]\} \text{ and } \{i : \mathscr{V}(\mathbb{N}) \models \sigma_j[\phi^1(i), \dots, \phi^n(i)]\} \in \mathfrak{U} \iff \\ &\mathscr{V}(\mathscr{N}) \models \sigma_i[\boldsymbol{\rho}(\phi^1), \dots, \boldsymbol{\rho}(\phi^n)] \text{ and } \mathscr{V}(\mathscr{N}) \models \sigma_j[\boldsymbol{\rho}(\phi^1), \dots, \boldsymbol{\rho}(\phi^n)] \iff \\ &\mathscr{V}(\mathscr{N}) \models \sigma_i \wedge \sigma_j[\boldsymbol{\rho}(\phi^1), \dots, \boldsymbol{\rho}(\phi^n)] \end{aligned}$$

This time it follows by the superset property of $\mathfrak{U}$. Note that the parameters $\phi^1, \dots, \phi^n$ are the same for σ_i, σ_j for readability, this is without loss of generalization.

For the existential quantifier, assume $(\star)$ holds for $\sigma[\xi_1, \dots, \xi_n]$, and suppose

$$D = \{i : \mathscr{V}(\mathbb{N}) \models \big(\exists x \in \phi^1(i)\big)\, \sigma[x, \phi^2(i), \dots, \phi^n(i)]\} \in \mathfrak{U}$$

Then define $\psi \in \mathscr{F}_\infty$ as follows:

$$\psi(i) = \begin{cases} x \in \phi^1(i) \text{ such that } \sigma[x, \phi^2(i), \dots, \phi^n(i)] \text{ if } i \in D \\ 1 \text{ otherwise} \end{cases}$$

Then $\{i : \mathscr{V}(\mathbb{N}) \vDash \sigma[\psi(i), \phi^2(i), \dots, \phi^n(i)]\} \in \mathfrak{U}$ so by assumption $\mathscr{V}(\mathscr{N}) \vDash \sigma[\boldsymbol{\rho}(\psi), \dots, \boldsymbol{\rho}(\phi^n)]$. Notice that $\psi \in_{\mathfrak{U}} \phi^1$, so $\boldsymbol{\rho}(\psi) \in \boldsymbol{\rho}(\phi^1)$ by proposition 3.4.6, hence $\mathscr{V}(\mathscr{N}) \vDash \left(\exists x \in \boldsymbol{\rho}(\phi^1)\right) \sigma[x, \boldsymbol{\rho}(\phi^2), \dots, \boldsymbol{\rho}(\phi^n)]$, so "$\Rightarrow$" holds.

For "$\Leftarrow$", assume $\mathscr{V}(\mathscr{N}) \vDash \left(\exists x \in \boldsymbol{\rho}(\phi^1)\right) \sigma[x, \boldsymbol{\rho}(\phi^2), \dots, \boldsymbol{\rho}(\phi^n)]$. Then there exists $\boldsymbol{\rho}(\psi) \in \boldsymbol{\rho}(\phi^1)$ such that $\mathscr{V}(\mathscr{N}) \vDash \sigma[\boldsymbol{\rho}(\psi), \dots, \boldsymbol{\rho}(\phi^n)]$ (note the quantification is not over urelements, so this follows by the definition in proposition 3.4.6). Then by assumption $\{i : \mathscr{V}(\mathbb{N}) \vDash \sigma[\psi(i), \phi^2(i), \dots, \phi^n(i)]\} \in \mathfrak{U}$. Also $\psi \in_{\mathfrak{U}} \phi^1$, so $\{i : \psi(i) \in \phi^1(i)\} \in \mathfrak{U}$, and thus:

$$\begin{aligned} \mathfrak{U} \ni \; & \{i : \psi(i) \in \phi^1(i)\} \cap \{i : \mathscr{V}(\mathbb{N}) \vDash \sigma[\psi(i), \phi^2(i), \dots, \phi^n(i)]\} \\ & = \{i : \mathscr{V}(\mathbb{N}) \vDash \left(\exists x \in \phi^1(i)\right) \sigma[x, \phi^2(i), \dots, \phi^n(i)]\} \end{aligned}$$

Which completes the proof. Closeness under equivalence of $L_\in$ formulas is a trivial step in the proof, so by induction $(\star)$ holds for all bounded $L_\in$ formulas.

$\square$

The proof of the first part of the theorem now follows easily, for $\sigma[\xi_1, \dots, \xi_n]$ an $L_\in$ formula, $x_1, \dots, x_n \in \mathscr{V}(\mathbb{N})$:

$$\begin{aligned} & \mathscr{V}(\mathbb{N}) \vDash \sigma[x_1, \dots, x_n] \iff \mathscr{V}(\mathbb{N}) \vDash \sigma[{}^c x_1(i), \dots, {}^c x_n(i)], i \in \mathbb{N} \iff \\ & \{i : \mathscr{V}(\mathbb{N}) \vDash \sigma[{}^c x_1(i), \dots, {}^c x_n(i)]\} \in \mathfrak{U} \iff \mathscr{V}(\mathscr{N}) \vDash \sigma[\boldsymbol{\rho}({}^c x_1), \dots, \boldsymbol{\rho}({}^c x_n)] \iff \\ & \mathscr{V}(\mathscr{N}) \vDash \sigma[*x_1, \dots, *x_n] \end{aligned}$$

The second property of a nonstandard embedding follows as well: ${}^*\mathbb{N} = \boldsymbol{\rho}({}^c\mathbb{N}) = \{\rho(\phi) : \phi \in \mathscr{F}\} = \{\mathfrak{n}(\mathbf{A}_\phi) : \phi \in \mathscr{F}\} = \mathscr{N}$. The third property holds as well, let A be an entity of $\mathscr{V}(\mathbb{N})$, then:

$${}^*A = \{\boldsymbol{\rho}(\phi) : \phi \in_{\mathfrak{U}} {}^c A\} \supset \{\boldsymbol{\rho}({}^c a) : a \in A\} = \{{}^*a : a \in A\}$$

So indeed $\langle \mathscr{V}(\mathbb{N}), \mathscr{V}(\mathscr{N}), *\rangle$ is a nonstandard universe. The only thing left to prove is that $*$ is ω-saturated. Therefore, consider a countable family of internal sets $\mathscr{B}$ that satisfies the finite intersection property. Observe that all elements of this set are internal, so by the structure of $\mathscr{V}(\mathbb{N})$ there is a set A such that $\mathscr{B} \subset {}^*A$. Therefore, the elements of $\mathscr{B}$ can be written as $\boldsymbol{\rho}(\phi^i)$ for $\phi^i : \mathbb{N} \to A, i \in \mathbb{N}$.

By the finite intersection property, for each $n \in \mathbb{N}$, $\cap_{i=0}^n \boldsymbol{\rho}(\phi^i) \neq \varnothing$. So for each n there exists $x_n \in \cap_{i=0}^n \boldsymbol{\rho}(\phi^i)$. The elements of $\mathscr{B}$ are internal sets, so x_n is by definition of the form $\boldsymbol{\rho}(\psi^n)$ for some $\psi^n \in \mathscr{F}_\infty$. It may be assumed that $\psi^n(k) \in \phi^i(k)$ for all $k \in \mathbb{N}, i \leq n$, the image under $\boldsymbol{\rho}$ does not differ. With this, define $\theta \in \mathscr{F}_\infty$ as $\theta(n) = \psi^n(n)$. Then for all $i \in \mathbb{N}$:

$$\{k : \theta(k) \in \phi^i(k)\} = \{k : \psi^k(k) \in \phi^i(k)\} = \{k : k \geq i\} \in \mathfrak{U} \text{ (it is cofinite)}$$

So $\theta \in_{\mathfrak{U}} \phi^i$ for all i. It follows $\boldsymbol{\rho}(\theta) \in \cap\mathscr{B} \neq \varnothing$, which was all to prove.

$\square$

3.5 Existence of a numerosity function on labelled sets

This section is devoted to proving the existence of a numerosity function by assuming a *Ramsey ultrafilter*[1]. This should be defined first:

Definition 3.5.1. *An ultrafilter* $\mathfrak{U}$ *on* $\mathbb{N}$ *is called **Ramsey** if for every partition* $\{X_n : n \in \mathbb{N}\}$ *of* $\mathbb{N}$ *such that* $X_n \notin \mathfrak{U}$ *for all* n*, there exists* $D \in \mathfrak{U}$ *such that* $D \cap X_n$ *contains precisely one element for each* n*.*

The name Ramsey comes from *Ramsey's theorem* in combinatorics about colouring of complete graphs. For ultrafilters, something similar holds, which is stated in the following proposition. It will not be proven here, but the proof can be found in [4], lemma 9.2, page 108. For any set $A \subseteq \mathbb{N}$, let $[A]^n$ denote the set of subsets of A with n elements.

Proposition 3.5.2. *A nonprincipal ultrafilter* $\mathfrak{U}$ *on* $\mathbb{N}$ *is Ramsey if and only if for any finite partition* $\{Y_1, \dots, Y_k\}$ *of* $[\mathbb{N}]^n$ *(*$n \in \mathbb{N}$ *arbitrary), there is a set* $H \in \mathfrak{U}$ *such that* $[H]^n \subseteq Y_i$ *for some* $i \in \{1, \dots, k\}$*. This* H *is also called a homogeneous set.*

Notice the partitioning can be thought of as "colouring" $[\mathbb{N}]^n$.

The following proposition should remind the reader of proposition 3.4.4:

Proposition 3.5.3. *A nonprincipal ultrafilter* $\mathfrak{U}$ *on* $\mathbb{N}$ *is Ramsey if and only if for any* $\phi \in \mathbb{N}^{\mathbb{N}}$ *there exists* $D \in \mathfrak{U}$ *such that* $\phi|_D$ *is nondecreasing (equivalently,* ϕ *is* $\mathfrak{U}$*-equivalent to a nondecreasing* ψ*).*

Proof. "$\Rightarrow$": Assume $\phi \in \mathbb{N}^{\mathbb{N}}$ arbitrary. Consider the partition $X_n = \{i : \phi_i = n\}$. Then if for some n, $X_n \in \mathfrak{U}$, then $\phi|_{X_n}$ is nondecreasing so we're done.

Suppose $X_n \notin \mathfrak{U}$ for all n. Then by assumption, there exists $D \in \mathfrak{U}$ such that $|D \cap X_n| = 1$. It follows $\phi|_D$ is injective. Define the following sequence inductively:

$$n_0 = \min D; n_{k+1} = \max\{n \in D : \phi(n) \leq \phi(m) \text{ for some } m \leq \xi_k\}$$

Where $\xi_k = \min\{n \in D : n > n_k\}$. Notice each set over which is taken the minimum or the maximum is finite, as $\phi|_D$ is injective. The sequence is increasing: $n_{k+1} \geq \xi_k > n_k$. Now consider the partition $\{Y_i : i \in \mathbb{N}\}$ defined by:

$$Y_0 = \{0, \dots, n_0\}; Y_{k+1} = \{n_k + 1, \dots, n_{k+1}\}$$

Then again there exists a set $E \in \mathfrak{U}$ such that $|E \cap Y_i| = 1$ for each i. By the properties of an ultrafilter, for $Y_e = \bigcup_{k \in \mathbb{N}} Y_{2k}$, $Y_o = \bigcup_{k \in \mathbb{N}} Y_{2k+1}$, it holds $Y_e \in \mathfrak{U}$ or $Y_o \in \mathfrak{U}$. Assume $Y_e \in \mathfrak{U}$. Then let $F = E \cap Y_e \cap D \in \mathfrak{U}$. Then $\phi|_F$ is increasing: let $x < y$. Then $x \in Y_{2k}, y \in Y_{2l}$ for some $k < l$. Then $y > n_{2k+1}$ (as h is at least $k+1$), so $\phi(y) > \phi(m)$ for all $m \leq \xi_{2k} \leq n_{2k+1}$ by definition of n_{2k+1}. So in particular $x \leq n_{2k+1}$, so $\phi(x) < \phi(y)$. The case where $Y_0 \in \mathfrak{U}$ is the same. Hence $\phi|_F$ is nondecreasing.

---

[1] In [1] this is called a *selective* ultrafilter. In [4], page 76, and [6], chapter VI, it is called Ramsey, and I decided to adopt that to be more in line with the common literature.

"$\Leftarrow$" Assume $\{X_n : n \in \mathbb{N}\}$ is a partition with $X_n \notin \mathfrak{U}$ for all n. Define $\phi \in \mathbb{N}^{\mathbb{N}}$ as $\phi(m) = n \iff m \in X_n$. By assumption, there is a $D \in \mathfrak{U}$ such that ϕ_D is nondecreasing. Define $D_n = D \cap X_n$ for any n. Each D_n is finite:

Suppose D_n is infinite for some n. Then for any $m > n$, there is no $k > \min D_n$ such that $\phi(k) = m$ (if there was, ϕ_D would be decreasing for some $l > k$). But then D_n is cofinite, as it contains all $k \geq \min D_n$, so $D_n \subseteq X_n \in \mathfrak{U}$, a contradiction.

So each D_n is finite, say $D_n = \{m_1^{(n)} < \cdots < m_{k_n}^{(n)}\}$.

With this notation, define $\psi \in \mathbb{N}^{\mathbb{N}}$ as $\psi(m_i^{(n)}) = k_n - i, \psi(n) = 0$ if $n \notin D$. Remark that on any D_n, ψ is decreasing. However, there exists $D' \in \mathfrak{U}$ such that $\psi_{D'}$ is nondecreasing. Combining this gives that for any n, $D_n \cap D'$ contains at most one element. So if we let $E = D \cap D'$, then $X_n \cap E$ contains at most one element for each n. Pick arbitrary elements for any n with $X_n \cap E = \varnothing$, and construct $E' \supseteq E$ such that $|X_n \cap E'| = 1$ for every n. This completes the proof. ☐

Combining propositions 3.4.4 and 3.5.3 gives that by assuming a numerosity function, one obtains a Ramsey ultrafilter. This is then enough to produce a nonstandard model $\mathscr{N}$ of the natural numbers, the set of numerosities. In looking for a foundation for the existence of a numerosity function, it seems natural to reverse the direction, and assume a Ramsey ultrafilter instead. This indeed, results in the following theorem.

Theorem 3.5.4 (Benci-Di Nasso)**.** *There exists a numerosity function if and only if there exists a Ramsey ultrafilter on* $\mathbb{N}$*.*

Proof. "$\Rightarrow$": Proven by combining propositions 3.4.4 and 3.5.3. "$\Leftarrow$": Let $\mathfrak{U}$ be a Ramsey ultrafilter on $\mathbb{N}$. Define $\mathscr{N}$ as the ultraproduct $\mathscr{N} = \mathbb{N}^{\mathbb{N}}/\mathfrak{U}$. Then the induced order $[\phi]_{\mathfrak{U}} \leq [\psi]_{\mathfrak{U}} \iff \phi \leq_{\mathfrak{U}} \psi \iff \{i : \phi_i \leq \psi_i\} \in \mathfrak{U}$ is a linear order on $\mathscr{N}$. For a labelled set $\mathbf{A}$, define $\mathfrak{n}(\mathbf{A}) = [\gamma_A]_{\mathfrak{U}}$, where $\gamma_A : n \mapsto |A_n|$ is the approximating sequence of $\mathbf{A}$. As $\mathfrak{U}$ is Ramsey, any $\phi \in \mathbb{N}^{\mathbb{N}}$ is $\mathfrak{U}$-equal to a nondecreasing one (an approximating sequence), so $\mathfrak{n}$ is surjective. What's left to check is properties (i)-(iii) of definition 3.1.3.

Property (i) follows immediately from the induced ordering: if $|A_n| \leq |B_n|$, then $\mathfrak{n}(\mathbf{A}) = [\gamma_A]_{\mathfrak{U}} \leq [\gamma_B]_{\mathfrak{U}} = \mathfrak{n}(\mathbf{B})$.

Recall property (ii): $\xi < \mathfrak{n}(\mathbf{A}) \iff \exists \mathbf{B} \subset \mathbf{A}$ such that $\xi = \mathfrak{n}(\mathbf{B})$. First assume labelled sets $\mathbf{B} \subset \mathbf{A}$. For an element $a \in A \setminus B$, let $m = \ell_A(a)$. Then for all $n \geq m$, $|B_n| < |A_n|$, hence $\{i : \gamma_B(i) < \gamma_A(i)\}$ is cofinite, thus in $\mathfrak{U}$. Therefore $\mathfrak{n}(\mathbf{B}) < \mathfrak{n}(\mathbf{A})$.

For the converse, assume $\xi = [\phi]_{\mathfrak{U}} < \mathfrak{n}(\mathbf{A}) = [\gamma_A]_{\mathfrak{U}}$. From ϕ, $\mathbf{B} \subset \mathbf{A}$ will be constructed to satisfy $\mathfrak{n}(\mathbf{B}) = \xi$.

The claim is that there exists a set $H = \{k_0 < k_1 < \dots\} \in \mathfrak{U}$ such that:

$$\phi(k_n) - \phi(k_{n-1}) \leq \gamma_A(k_n) - \gamma_A(k_{n-1}) \tag{3.1}$$

This follows from proposition 3.5.2. Consider $Y = \{\{m, m'\} \in [\mathbb{N}]^2 : \phi(m) - \phi(m') \leq \gamma_A(m) - \gamma_A(m')$ where w.l.o.g. $m < m'\}$, $Y' = [\mathbb{N}]^2 \setminus Y$, and let $\{Y, Y'\}$ be the corresponding partition of $[\mathbb{N}]^2$. By the proposition there exists a homogeneous $H = \{k_0 < k_1 < \dots\} \in \mathfrak{U}$ for this partition. But $[H]^2 \subseteq Y'$ cannot be, for then it must hold for two consecutive $k_{n-1} < k_n$:

$$\begin{aligned} &\phi(k_n) - \phi(k_{n-1}) > \gamma_A(k_n) - \gamma_A(k_{n-1}) \iff \\ &\phi(k_n) - \phi(k_{n-1}) \geq \gamma_A(k_n) - \gamma_A(k_{n-1}) + 1 \iff \\ &\phi(k_n) - \phi(k_0) \geq \gamma_A(k_n) - \gamma_A(k_0) + n \text{ for all } n \end{aligned}$$

So then it holds for all $m > \phi(k_0) - \gamma_A(k_0)$ that $\phi(k_m) > \gamma_A(k_m)$, which implies $\{i : \phi(i) > \gamma_A(i)\} \supseteq \{k_m, k_{m+1}, \dots\} \in \mathfrak{U}$ which is contradictory. Hence $[H]^2 \subseteq Y$.

Now let $H' = H \cap \{i : \phi(i) < \gamma_A(i)\} \in \mathfrak{U}$, write $H' = \{k'_0 < k'_1 < k'_2 < \dots\}$. Define $B_0 \subseteq \{a \in A : \ell_A(a) \leq k'_0\}$ such that $|B_0| = \phi(k_0)$. Proceed by induction: let $B_n \subseteq \{a \in A : k'_{n-1} < \ell_A(a) \leq k'_n\}$ with $|B_n| = \phi(k'_n) - \phi(k'_{n-1})$. This can be done because of (3.1). The labelled set $\mathbf{B} \subset \mathbf{A}$ is now obtained by letting $B = \bigcup_i B_i$, with induced labelling function γ_B. For $h \in H'$ we have $\phi(h) = \gamma_B(h)$, so $\gamma_B =_{\mathfrak{U}} \phi$. Hence $\mathfrak{n}(\mathbf{B}) = [\gamma_B]_{\mathfrak{U}} = [\phi]_{\mathfrak{U}} = \xi$.

At last, the third property follows from the fact that $\mathfrak{n}(\mathbf{A}) = \mathfrak{n}(\mathbf{A}')$ and $\mathfrak{n}(\mathbf{B}) = \mathfrak{n}(\mathbf{B}')$ gives $\gamma_A =_{\mathfrak{U}} \gamma_{A'}$ and $\gamma_B =_{\mathfrak{U}} \gamma_{B'}$. Because then $\gamma_{A \oplus B} =_{\mathfrak{U}} \gamma_A + \gamma_B =_{\mathfrak{U}} \gamma_{A'} + \gamma_{B'} =_{\mathfrak{U}} \gamma_{A' \oplus B'}$, and $\gamma_{A \odot B} =_{\mathfrak{U}} \gamma_A \cdot \gamma_B =_{\mathfrak{U}} \gamma_{A'} \cdot \gamma_{B'} =_{\mathfrak{U}} \gamma_{A' \odot B'}$. ☐

Another result stated in [4], but attributed to W. Rudin is that the continuum hypothesis implies the existence of a Ramsey ultrafilter. Also, a result by S. Shelah (see [6], chapter VI) is that the absence of Ramsey ultrafilter is consistent with ZFC. So, as CH is independent of ZFC, one can formulate the following corollary:

Corollary 3.5.5. *The existence of a numerosity function on labelled sets is independent of ZFC.*

This result is of course about the numerosity function defined in the way that it is done here.

Chapter 4

Numerosities of sets of ordinals

In this chapter, the idea of a numerosity is slightly modified in order to make a generalisation to arbitrary sets of ordinals. Uncountable sets cannot be treated using the method of labelled sets as easy as countable sets can. Ordinals, however, provide some tools that can be used to make a similar construction possible. In the following, this is done mostly in analogy with [2]. However, the definition of a numerosity function as is done there gives trouble in proving a theorem that is claimed proven there. I solved that issue by slightly modifying the original definition. At the end, I compare the numerosity function of chapter 3 with the new one that provides a new insight in how the two definitions might be unified.

4.1 Defining the numerosity function

The goal is to define a numerosity function on sets of ordinals, i.e. a function $\mathfrak{n}$ with domain $\mathcal{W} = \mathscr{P}(Ord)$. The idea of labelling is dropped here, instead $\mathfrak{n}$ will be defined explicitly for these sets. As these sets are all sets of ordinals, the labelling can be thought of as "built in", like the canonical labelling of $\mathbb{N}$ in the previous section.

Recall that the motivation behind a numerosity is a more natural, "aristotelian" notion of size. The numerosities should therefore be elements of a ring structure, to have addition and multiplication, endowed with an ordering that will provide the notion of size. Such a structure $\mathcal{A}$ might even be a proper class to provide enough numerosities for every $A \in \mathcal{W}$, but we could always restrict $\mathfrak{n}$ to elements of $\mathcal{W}_\kappa = \mathscr{P}(\kappa)$ for some cardinal κ in order to deal with sets only. However, in appendix A.2 we justify dealing with proper classes in the rest of this text by assuming an inaccessible cardinal.

For a numerosity function, a natural arithmetic is desirable. In the case of ordinals, this is not easily provided, as ordinals have strange arithmetic themselves. For example, it cannot be guaranteed that the following holds for all ordinals τ:

$$\mathfrak{n}(A) = \mathfrak{n}(\{\tau\} + A) = \mathfrak{n}(\{\tau + \alpha | \alpha \in A\})$$
$$\mathfrak{n}(A) = \mathfrak{n}(\{\tau\} \cdot A) = \mathfrak{n}(\{\tau \cdot \alpha | \alpha \in A\})$$

This property is however desirable, because the "size" of a set of ordinals should be ideally independent of the ordinals it contains, and the identities above are special cases of this property: translation and multiplication invariance.

Here the set operations + and $\cdot$ are as usual when the elements are part of a (semi-)ring, so they're applied to all elements of both sets. These properties, that will be called the *translation* and *multiplication* invariance of the numerosity function, cannot be satisfied in general. It will however be satisfied for τ sufficiently large compared to A, this will be made precise shortly.

An ordinal of the form ω^α, with $\alpha > 0$, will be called a *tile* here, in analogy with [2]. An important property of a tile $\tau = \omega^\alpha$ is that it can be written as a product of atom powers by expanding α in its normal form. Let $\alpha = \omega^{\alpha_1} a_1 + \dots + \omega^{\alpha_n} a_n$. Then for $\theta_i = \omega^{\omega^{\alpha_i}}$:

$$\tau = \theta_1^{a_1} \cdot \dots \cdot \theta_n^{a_n}$$

This product representation is unique according to the Cantor normal form theorem.

One can also express τ as power θ^γ: Let $\theta = \theta_n$, and let δ_i be such that $\alpha_i = \alpha_n + \delta_i$. Then for $\gamma = \omega^{\delta_1} a_1 + \dots + \omega^{\delta_{n-1}} a_{n-1} + a_n$, $\tau = \theta^\gamma$:

$$\theta^\gamma = \omega^{\omega^{\alpha_n} \omega^{\delta_1} a_1 + \dots + \omega^{\delta_{n-1}} a_{n-1} + a_n} = \omega^{\omega^{\alpha_1} a_1 + \dots + \omega^{\alpha_n} a_n} = \tau$$

This will be an important (unique) characterization of τ, and in the following τ is called a θ-tile if it is a power of an atom θ, following the terminology in [2].

In order to provide a certain product principle like there was in the countable case (the cartesian product principle), a new product operation on sets of ordinals must be defined. Notice that a cartesian product of two such sets is no longer a set of ordinals, so to have some product principle of a numerosity function on $\mathcal{W}$, there should a product on $\mathcal{W}$ in the first place. For τ an infinite tile, define the product $\otimes_\tau : \mathcal{W} \times \mathcal{W} \to \mathcal{W}$ as:

$$A \otimes_\tau B = \{\tau\beta + \alpha | \alpha \in A, \beta \in B\}$$

This *flat product* on $\mathcal{W}$ will be a very useful definition. For $\tau = \omega^\alpha$ a θ-tile, notice that in the product $\tau \otimes_\tau \delta$ there will be no absorption if $\delta \leq \theta^\omega$: For $\beta < \delta$, we then obtain $\beta < \theta^\omega = \left(\omega^{\omega^{\alpha_n}}\right)^\omega = \omega^{\omega^{\alpha_n + 1}}$. So β is always of the form $\beta = \omega^{\omega^{\alpha_n} k} b_1 + \omega^{\beta_2} b_2 + \dots + \omega^{\beta_m} b_m$ for finite k, and $\omega^{\alpha_n} k > \beta_2 > \dots > \beta_m$, b_i finite with b_1 possibly zero. Then any element of $\tau \otimes_\tau \delta$ is of the form $\tau\beta + \alpha$ with $\alpha < \tau$ and β as above:

$$\begin{aligned} \tau\beta + \alpha \quad &= \omega^{\omega^{\alpha_1} a_1 + \dots + \omega^{\alpha_n} a_n} \left(\omega^{\omega^{\alpha_n} k} b_1 + \omega^{\beta_2} b_2 + \dots + \omega^{\beta_m} b_m\right) + \alpha \\ &= \omega^{\omega^{\alpha_1} a_1 + \dots + \omega^{\alpha_n} (a_n + k)} b_1 + \textstyle\sum_{i=2}^m \omega^{\omega^{\alpha_1} a_1 + \dots + \omega^{\alpha_n} a_n + \beta_i} b_i + \alpha \end{aligned}$$

So no absorption occurs. Note that if $\beta = \theta^\omega$ it goes wrong, so $\delta = \theta^\omega$ is the largest δ we can demand for this flat product to "behave like a cartesian product", that is: an element in $\tau \otimes_\tau \delta$ corresponds to a unique pair (α, β).

So if for certain sets $A, B \in \mathcal{W}$, $A \otimes_\tau B$ "behaves like a cartesian product", it seems valuable to ask the numerosity function to satisfy the following:

$$\mathfrak{n}(A \otimes_\tau B) = \mathfrak{n}(A) \cdot \mathfrak{n}(B)$$

This is called the *flat product principle* in [2]. In the definition below, I'll just refer to it as *product principle.*

For ordinals, $\otimes_\tau$ can be used to represent division with remainder. Suppose α is an ordinal, and let $\alpha = \tau\beta + \gamma$ be the representation after division by τ, so $\gamma < \tau$. Then $\gamma \otimes_\tau \beta = \{\tau\lambda + \mu | \lambda < \beta, \mu < \gamma\} = \{\kappa | \kappa < \alpha\} = \alpha$. So division by τ can be written as $\alpha = \gamma \otimes_\tau \beta$.

The above discussion provides enough machinery to state the numerosity function $\mathfrak{n}$ and show some properties.

Definition 4.1.1. *For $\mathcal{W} = \mathscr{P}(Ord)$ the class of all sets of ordinals, $\mathcal{A}$ an ordered class-ring, a numerosity function $\mathfrak{n} : \mathcal{W} \to \mathcal{A}$ takes positive values in $\mathcal{A}$ and satisfies the following properties:*

(i) Cantor principle: If $|X| < |Y|$, then $\mathfrak{n}(X) < \mathfrak{n}(Y)$

(ii) Sum principle: $X \cap Y = \varnothing$, then $\mathfrak{n}(X \cup Y) = \mathfrak{n}(X) + \mathfrak{n}(Y)$.

(iii) Product principle: For τ a tile with base θ, $\mathfrak{n}(X) \cdot \mathfrak{n}(Y) = \mathfrak{n}(X \otimes_\tau Y)$ for all $X \subset \tau$, $Y \subseteq \theta^\omega$.

(iv) Unit principle: For every ordinal α, $\mathfrak{n}(\{\alpha\}) = 1$.

In [2], the Cantor principle is stated as: $\mathfrak{n}(X) = \mathfrak{n}(Y) \Rightarrow |X| = |Y|$. That seems like a weaker statement than the above, and that creates trouble in proving a necessary and sufficient condition for this property later on (theorem 4.2.4). Stating it this way solves the problem, and it is obvious that we may demand such a property from a numerosity function. Notice as well that I name it a Cantor principle, although it is actually a half-Cantor principle. It is however redundant to say that, as obviously a numerosity does not coincide with cardinality.

Also, in [2] the requirement for Y in property (iii) is stated as $Y \subseteq \delta < \theta^\omega$ for some δ. As I've shown above in the discussion of the flat product, this range can be extended and simplified to $Y \subseteq \theta^\omega$, which is what I've done here. It won't lead to any trouble later on.

A few properties that follow from the definition are immediate. For instance, if $A \subset B \in \mathscr{P}(Ord)$, let $\beta \in B \setminus A$ be arbitrary. Then the aristotelian property is easily verified:

$$\mathfrak{n}(B) = \mathfrak{n}(A) + \mathfrak{n}(\{\beta\}) + \mathfrak{n}((B \setminus A) \setminus \{\beta\}) \geq \mathfrak{n}(A) + 1 > \mathfrak{n}(A)$$

Also, for finite sets $\mathfrak{n}(\{\alpha_1, \dots, \alpha_n\}) = \mathfrak{n}(\{\alpha_1\}) + \dots + \mathfrak{n}(\{\alpha_n\}) = n = |\{\alpha_1, \dots, \alpha_n\}|$ if the subsemi-ring generated by 1 of $\mathcal{A}$ (i.e. the set of elements of the form $1 + \dots + 1$) is identified with $\mathbb{N}$. So for finite set, this notion of size agrees with cardinality.

The following proposition proves the translational and multiplicative invariance:

Proposition 4.1.2. *For $\mathfrak{n}$ as in definition 4.1.1, τ a tile the following holds:*

(i) If $A \subseteq \tau \cdot n$ for some finite n, then $\mathfrak{n}(A) = \mathfrak{n}(\{\tau \cdot \beta\} + A)$ for all ordinals β.

(ii) For τ a tile with atom base θ, $\mathfrak{n}(A) = \mathfrak{n}(\{\tau\} \cdot A)$ for all $A \subseteq \theta^\omega$.

Proof. (i) The proof is by induction on β. This is best done by first proving it for $\beta = m < \omega$, and then for $\beta > \omega$ all at once.

So first assume $\beta = m < \omega$. Remark that because $A \subseteq \tau \cdot n$, with τ a tile, considering the normal form of an $\alpha \in A$ yields $\alpha = \tau \cdot i + \delta$ where $\delta < \tau$, and $0 \leq i \leq n-1$. Hence A is the disjoint union of n sets: $A = \bigcup_{i=0}^{n-1} \{\tau \cdot i\} + A_i$ with $A_i \subseteq \tau$. Then:

$$\begin{aligned}
\mathfrak{n}\left(\{\tau \cdot m\} + A\right) &= \mathfrak{n}\left(\{\tau \cdot m\} + \left(\textstyle\bigcup_{i=0}^{n-1} \{\tau \cdot i\} + A_i\}\right)\right) = \mathfrak{n}\left(\textstyle\bigcup_{i=0}^{n-1} \{\tau \cdot (i+m)\} + A_i\}\right) \\
&= \textstyle\bigcup_{i=0}^{n-1} \mathfrak{n}\left(\{\tau \cdot (i+m)\} + A_i\right) = \bigcup_{i=1}^{n} \mathfrak{n}\left(A_i \otimes_\tau \{i+m\}\right) \\
&= \textstyle\bigcup_{i=0}^{n} \mathfrak{n}\left(A_i\right) \cdot 1 = \mathfrak{n}\left(\bigcup_{i=0}^{n} A_i\right) \\
&= \mathfrak{n}\left(A\right)
\end{aligned}$$

Note each step is allowed, as the union is a disjoint union each time, and we make use of the sum, product and unit principles. The product principle is allowed as $A_i \subset \tau$, and $\{i+m\}$ is finite.

For the other case, it may be assumed that β is not a tile, for else $\tau' = \tau \cdot \beta$ is a tile, and the proposition just holds for τ': $\mathfrak{n}\left(\{\tau'\} + A\right) = \mathfrak{n}\left(A\right)$.

So assume $\beta = \omega^{\beta_1}(m+1) + \delta$ (in normal form), for $0 \leq m < \omega$. Trivially $\omega^{\beta_1} m + \delta < \beta$, so as induction hypothesis the proposition is assumed for $\omega^{\beta_1} m + \delta$. Define the tile $\tau' = \tau \cdot \omega^{\beta_1}$. Then:

$$\begin{aligned}
\mathfrak{n}\left(\{\tau \cdot \beta\} + A\right) &= \mathfrak{n}\left(\{\tau'\} + \left(\{\omega^{\beta_1} m + \delta\} + A\right)\right) \\
&=^! \mathfrak{n}\left(\{\omega^{\beta_1} m + \delta\} + A\right) \\
&= \mathfrak{n}\left(A\right)
\end{aligned}$$

Where $=^!$ is valid because $\left(\{\omega^{\beta_1} m + \delta\} + A\right) \subseteq (\beta + \tau) \subseteq \tau'$ so the proposition is applied for the already proven finite case.

The proof now follows by induction on β.

(ii) Note $\{\tau\} \cdot A = \{0\} \otimes_\tau A$. So then $\mathfrak{n}\left(\{\tau\} \cdot A\right) = \mathfrak{n}\left(\{0\} \otimes_\tau A\right) = 1 \cdot \mathfrak{n}\left(A\right)$ follows immediately by the product principle, under the given conditions. □

Another property of the numerosity function on sets of ordinals is that it commutes with Cantor's normal form of an ordinal, and also with the product expansion. This is stated in the following proposition:

Proposition 4.1.3. *Let* α *be an ordinal, with normal form* $\alpha = \sum_{i=1}^{n} \omega^{\alpha_i} a_i$ *, and* τ *a tile with product expansion* $\tau = \prod_{i=1}^{n} \theta_i^{a_i}$

Then:

(i)

$$\mathfrak{n}\left(\sum_{i=1}^{n} \omega^{\alpha_i} a_i\right) = \sum_{i=1}^{n} \mathfrak{n}\left(\omega^{\alpha_i} a_i\right)$$

(ii)

$$\mathfrak{n}\left(\prod_{i=1}^{n} \theta_i^{a_i}\right) = \prod_{i=1}^{n} \mathfrak{n}\left(\theta_i^{a_i}\right)$$

Proof. (i) This follows easily by observing that $\{\omega^{\alpha_1}a_1\} + \sum_{i=2}^{n}\omega^{\alpha_i}a_i = \{\beta : \omega^{\alpha_1}a_1 \leq \beta < \alpha\}$, and $\omega^{\alpha_1}a_1 = \{\beta : \beta < \omega^{\alpha_1}a_1\}$, so that $\alpha = \{\beta : \beta < \alpha\} = (\omega^{\alpha_1}a_1) \cup (\{\omega^{\alpha_1}a_1\} + \sum_{i=2}^{n}\omega^{\alpha_i}a_i)$. The following now proves the statement:

$$\begin{aligned}
\mathfrak{n}\left(\textstyle\sum_{i=1}^{n}\omega^{\alpha_i}a_i\right) &= \mathfrak{n}\left((\omega^{\alpha_1}a_1) \cup (\{\omega^{\alpha_1}a_1\} + \textstyle\sum_{i=2}^{n}\omega^{\alpha_i}a_i)\right) \\
&= \mathfrak{n}\left(\omega^{\alpha_1}a_1\right) + \mathfrak{n}\left(\{\omega^{\alpha_1}a_1\} + \textstyle\sum_{i=2}^{n}\omega^{\alpha_i}a_i\right) \\
&=^{!} \mathfrak{n}\left(\omega^{\alpha_1}a_1\right) + \mathfrak{n}\left(\textstyle\sum_{i=2}^{n}\omega^{\alpha_i}a_i\right) \\
&= \textstyle\sum_{i=1}^{n}\mathfrak{n}\left(\omega^{\alpha_i}a_i\right)
\end{aligned}$$

The last identity is obtained by repeating the expansion $n-2$ times. At $=^{!}$ the previous proposition is applied, for $\tau = \omega^{\alpha_1}$ and using $\sum_{i=2}^{n}\omega^{\alpha_i}a_i \subset \omega^{\alpha_1}a_1$ (by the normal form theorem).

(ii) Recall that $\tau = \theta_1^{a_1} \otimes_{\theta_1^{a_1}} (\theta_2 \cdot \cdots \cdot \theta_n)$ by division of ordinals. $\theta_1^{a_1}$ Has base θ_1, and by definition $\theta_2^{a_1} \cdot \cdots \cdot \theta_n^{a_n} = \omega^{\omega^{\alpha_2}a_2 + \cdots + \omega^{\alpha_n}a_n} < \omega^{\omega^{\alpha_1+1}} = \theta_1^{\omega}$, so the product principle holds. Combining these facts yields the following expansion rule:

$$\begin{aligned}
\mathfrak{n}\left(\textstyle\prod_{i=1}^{n}\theta_i^{a_i}\right) &= \mathfrak{n}\left(\theta_1^{a_1} \otimes_{\theta_1^{a_1}} (\theta_2^{a_2} \cdot \cdots \cdot \theta_n^{a_n})\right) = \mathfrak{n}\left(\theta_1^{a_1}\right) \cdot \mathfrak{n}\left(\theta_2^{a_2} \cdot \cdots \cdot \theta_n^{a_n}\right) \\
&= \textstyle\prod_{i=1}^{n}\mathfrak{n}\left(\theta_i^{a_i}\right)
\end{aligned}$$

Where the latter holds by repeatedly expanding the product $n-2$ times more.

□

The above properties can be exploited further by defining a ring-like structure on *Ord*, using a so-called natural sum $\oplus$ and product $\odot$. These operations will have the pleasing property that they are commutative, and like the sum and product operations of labelled sets, they will proven to satisfy $\mathfrak{n}(\alpha \oplus \beta) = \mathfrak{n}(\alpha) + \mathfrak{n}(\beta)$, and $\mathfrak{n}(\alpha \odot \beta) = \mathfrak{n}(\alpha) \cdot \mathfrak{n}(\beta)$.

The operations $\alpha \oplus \beta$ and $\alpha \odot \beta$ are defined by considering the normal form of α, β as formal polynomials in ω, and then taking sum and products as usual. For notational reasons, we gather the exponents of α and β in one set of exponents $\gamma_1, \ldots, \gamma_n$, where the corresponding coefficient a_i or b_i might be 0.

$$\begin{aligned}
\alpha \oplus \beta &= \left(\textstyle\sum_{i=1}^{n}\omega^{\gamma_i}a_i\right) \oplus \left(\textstyle\sum_{i=1}^{n}\omega^{\gamma_i}b_i\right) \\
&= \textstyle\sum_{i=1}^{n}\omega^{\gamma_i}(a_i + b_i) \\
\alpha \odot \beta &= \left(\textstyle\sum_{i=1}^{n}\omega^{\gamma_i}a_i\right) \odot \left(\textstyle\sum_{i=1}^{n}\omega^{\gamma_i}b_i\right) \\
&= \textstyle\bigoplus_{i,j=1}^{n}\omega^{\gamma_i \oplus \gamma_j}a_i b_j
\end{aligned}$$

As $\mathfrak{n}\left(\omega^{\delta}d\right) = \mathfrak{n}\left(\omega^{\delta} \otimes_{\omega^{\delta}} d\right) = \mathfrak{n}\left(\omega^{\delta}\right) \cdot d$, it follows that numerosity function induces a semi-ring homomorphism between $\langle \mathit{Ord}, \oplus, \odot \rangle$ and $\langle \mathcal{A}, +, \cdot \rangle$.

4.2 Foundation of the numerosity function

As seen in the previous section, a numerosity function on sets of ordinals has very interesting properties, with only a few assumptions made for the function itself. In this section, the foundation of this function with its properties is explained in terms of *finite approximations*, which is very much like the treatment in the countable case. The difference is that here is often dealt with classes due to the fact that sets might be too small to work with.

What is needed first is a stricter notion of what is meant by a *finite approximation* in the general case. In the case of a set $X \in \mathcal{W} = \mathscr{P}(\mathit{Ord})$, it is certainly not always possible to have finite approximations X_i such that:

$$X = \bigcup_{i \in \mathbb{N}} X_i$$

This works for the countable case of labelled sets, but certainly not for much bigger sets of ordinals. What is needed here is a class $\mathcal{I}$ (to keep things simple and general, a class is assumed for now) that is *directed. Directed* means that $\mathcal{I}$ is endowed with a preorder $\leq$, with the property that any two elements have an upper bound. A preorder is the a partial order without the property of antisymmetry, i.e. there can be $i, j \in \mathcal{I}$ with $i \neq j$, $i \leq j$ and $i \geq j$. The upper bound property implies that for any $i, j \in \mathcal{I}$ there is also a $k \in \mathcal{I}$ such that $i \leq k$ and $j \leq k$. Note that every linearly ordered set is a directed set, so the following general definition for a finite approximation also holds for the countable case of labelled sets, where $\mathcal{I} = \mathbb{N}$.

Definition 4.2.1. *For a class of sets $\mathscr{C}$, $\langle \mathcal{I}, \leq \rangle$ a directed class, a map $\phi : \mathscr{C} \times \mathcal{I} \to \mathscr{C}$ is called a* ***finite approximation*** *if, for all $X, Y \in \mathscr{C}$ and $i, j \in \mathcal{I}$, the following holds:*

(i) $\phi(X, i) \subseteq X$ *is finite*

(ii) For all $x \in X$ there exists i such that $x \in \phi(X, i)$

(iii) If $i \leq j$, $\phi(X, i) \subseteq \phi(X, j)$

(iv) $\phi(X \cup Y, i) = \phi(X, i) \cup \phi(Y, i)$

Like before, $\phi(X, i)$ is denoted shortly by X_i. That way, *the counting function associated to X* is the map $\mathbb{N}^{\mathcal{I}} \ni \Phi(X) : i \mapsto |X_i|$. In the case of sets of ordinals, a finite approximation is desired for $\mathscr{C} = \mathcal{W}$.

The idea is now as follows. The counting functions $\Phi(X)$ belonging to each $X \in \mathcal{W}$ are elements of the ring $\mathbb{Z}^{\mathcal{I}}$. A numerosity function is obtained by defining, for a given prime ideal $\mathfrak{p}$, the quotient $\mathcal{A} = \mathbb{Z}^{\mathcal{I}}/\mathfrak{p}$ with projection map $\pi : \mathbb{Z}^{\mathcal{I}} \to \mathbb{Z}^{\mathcal{I}}/\mathfrak{p}$, and letting $\mathfrak{n}(X) = \pi(\Phi(X))$. This will be made precise in the following theorem, but first there are some remarks to make.

Assume for the moment that $\mathcal{I}$ is a set.[1] We'd like to show that every prime ideal $\mathfrak{p} < \mathbb{Z}^{\mathcal{I}}$ defines has an *associated ultrafilter* on $\mathcal{I}$. This is shown in the following lemma:

Lemma 4.2.2. *Let $\mathfrak{p} < \mathbb{Z}^{\mathcal{I}}$ be a prime ideal. Consider for an arbitrary subset $A \subset I$, the element $\epsilon_A \in \mathbb{Z}^{\mathcal{I}}$:*

$$\epsilon_A(i) = \begin{cases} 1 \text{ if } i \notin A \\ 0 \text{ if } i \in A \end{cases}$$

Then $\epsilon_A \in \mathfrak{p} \Rightarrow A \in \mathfrak{U}$ for every $A \subset I$ defines an ultrafilter $\mathfrak{U}$ on $\mathcal{I}$. Conversely, for $\mathfrak{U}$ an ultrafilter, $A \in \mathfrak{U} \Rightarrow \epsilon_A \in \mathfrak{p}$ defines a prime ideal in $\mathbb{Z}^{\mathcal{I}}$.

Furthermore, the ultrafilter $\mathfrak{U}$ is nonprincipal if and only if $\mathfrak{p}$ is a nonprincipal ideal

---

[1] This is made precise in appendix A.2

Proof. Notice $\epsilon_A \cdot \epsilon_{\mathcal{I}\setminus A} = 0 \in \mathfrak{p}$ for each prime ideal $\mathfrak{p}$, hence $\epsilon_A \in \mathfrak{p}$ or $\epsilon_{\mathcal{I}\setminus A} \in \mathfrak{p}$, but not both, as $\epsilon_A + \epsilon_{\mathcal{I}\setminus A} = 1$. Furthermore $1 = \epsilon_{\varnothing} \notin \mathfrak{p}$. Also, for $\epsilon_A, \epsilon_B \in \mathfrak{p}$, $\epsilon_{\mathcal{I}\setminus A}, \epsilon_{\mathcal{I}\setminus B} \notin \mathfrak{p}$, hence $\epsilon_{\mathcal{I}\setminus A} \cdot \epsilon_{\mathcal{I}\setminus B} = \epsilon_{\mathcal{I}\setminus A \cup \mathcal{I}\setminus B} \notin \mathfrak{p}$, thus $\epsilon_{A\cap B} = \epsilon_{\mathcal{I}\setminus(\mathcal{I}\setminus A \cup \mathcal{I}\setminus B)} \in \mathfrak{p}$. At last, if $\epsilon_A \in \mathfrak{p}$, and $A \subset B$, then $\epsilon_{A\setminus\mathcal{I}} \cdot \epsilon_B = 0 \in \mathfrak{p}$ so $\epsilon_B \in \mathfrak{p}$ by the above. So indeed, $\mathfrak{p}$ defines an ultrafilter $\mathfrak{U}$ on $\mathcal{I}$, by definition A.1.1.

Now suppose one is given an ultrafilter $\mathfrak{U}$ on $\mathcal{I}$. Then define $\mathfrak{p}$ as follows: $\phi = \psi \pmod{\mathfrak{p}} \iff \{i \in \mathcal{I} : \phi_i = \psi_i\} \in \mathfrak{U}$, or equivalently $\phi \in \mathfrak{p} \iff \{i \in \mathcal{I} : \phi_i = 0\} \in \mathfrak{U}$. For $\phi \cdot \psi \in \mathfrak{p}$, $\{i : \psi_i \cdot \phi_i = 0\} = \{i : \psi_i = 0\} \cup \{i : \phi_i = 0\} \in \mathfrak{U}$, so ϕ or $\psi \in \mathfrak{p}$, or otherwise $\{i : \psi_i \neq 0\} \cap \{i : \phi_i \neq 0\} \in \mathfrak{U}$ which is a contradiction. So $\mathfrak{p}$ is a prime ideal (the fact that it is an ideal follows by definition and elementary properties of ultrafilters). Notice $A \in \mathfrak{U} \Rightarrow \epsilon_A \in \mathfrak{p}$, and vice versa by the above, hence the claim follows.

For the last claim: suppose $\mathfrak{U}$ is principal. Then $A \in \mathfrak{U} \iff X_0 \subseteq A$ for a finite set X_0. Then $\phi \in \mathfrak{p} \iff \{i : \phi_i = 0\} \subseteq X_0 \iff \phi = \epsilon_{X_0} \cdot \phi$. So $\mathfrak{p}$ is generated by the element ϵ_{X_0}. The converse is obvious. ☐

Corollary 4.2.3. *$\mathcal{A} = \mathbb{Z}^{\mathcal{I}}/\mathfrak{p}$ is isomorphic to the ultrapower $\prod_{\mathfrak{U}} \mathbb{Z}^{\mathcal{I}}$ (see appendix A.1.1).*

Now the theorem can be stated:

Theorem 4.2.4. *For $\mathcal{W} = \mathscr{P}(Ord)$, $\phi : \mathcal{W} \times \mathcal{I} \to \mathcal{W}$ a finite approximation with associated counting function $\Phi : \mathcal{W} \to \mathbb{N}^{\mathcal{I}}$, $\mathfrak{p}$ a prime ideal of $\mathbb{Z}^{\mathcal{I}}$ and $\pi : \mathbb{Z}^{\mathcal{I}} \to \mathbb{Z}^{\mathcal{I}}/\mathfrak{p} = \mathcal{A}$ the projection map, and $\mathfrak{U}$ the ultrafilter associated to $\mathfrak{p}$. Then for $\mathfrak{n} = \pi \circ \Phi$, the conditions $(i) - (iv)$ of definition 4.1.1 can be reformulated in the following way:*

(i) holds by definition

(ii) $C_{XY} = \{i \in \mathcal{I} \| X_i| < |Y_i|\} \in \mathfrak{U}$ whenever $|X| < |Y|$

(iii) $P^{\tau}_{XY} = \{i \in \mathcal{I} \| X_i| \cdot |Y_i| = |(X \otimes_\tau Y)_i|\} \in \mathfrak{U}$ whenever τ is a θ-tile with $X \subseteq \tau$ and $Y \subseteq \theta^\omega$.

(iv) $C_{\varnothing\{\alpha\}} \in \mathfrak{U}$ for all ordinals α, with $C_{\varnothing\{\alpha\}}$ as above.

Proof. In the following proof, the induced order on $\mathbb{Z}^{\mathcal{I}}/\mathfrak{p}$ is used, namely that $\pi(\xi) < \pi(\eta) \iff \{i \in \mathcal{I} : \xi_i < \eta_i\} \in \mathfrak{U}$. Note that $\mathfrak{U}$ being an ultrafilter makes this a linear order.

(i) Assume $X, Y \in \mathcal{W}$ with $X \cap Y = \varnothing$. Notice that this implies $X_i \cap Y_i = \varnothing$ for all $i \in \mathcal{I}$, so $|X_i \cup Y_i| = |X_i| + |Y_i|$. The proof now follows, using property (iv) of the finite approximation function:

$$\begin{aligned} \mathfrak{n}(X \cup Y) &= \pi \circ \Phi(X \cup Y) = \pi(i \mapsto |\phi(X \cup Y, i)|) = \pi(i \mapsto |\phi(X, i) \cup \phi(Y, i)|) \\ &= \pi(i \mapsto |X_i \cup Y_i|) = \pi(i \mapsto |X_i|) + \pi(i \mapsto |Y_i|) = \pi(\Phi(X)) + \pi(\Phi(Y)) \\ &= \mathfrak{n}(X) + \mathfrak{n}(Y) \end{aligned}$$

(ii) "$\Leftarrow$": Holds obviously by the induced order, for $C_{XY} \in \mathfrak{U}$ gives $\pi(\Phi(X)) < \pi(\Phi(Y))$ so $\mathfrak{n}(X) < \mathfrak{n}(Y)$. "$\Rightarrow$": The same, here $|X| < |Y|$ implies $\pi(\Phi(X)) < \pi(\Phi(Y)) \iff \{i : |X_i| < |Y_i|\} \in \mathfrak{U}$.

(iii) This is obvious, as $|X_i| \cdot |Y_i| = (\Phi(X) \cdot \Phi(Y))_i$ so $P^{\tau}_{XY} \in \mathfrak{U} \iff \pi(\Phi(X) \cdot \Phi(Y)) = \pi(\Phi(X \otimes_\tau Y)) \iff \mathfrak{n}(X) \cdot \mathfrak{n}(Y) = \mathfrak{n}(X \otimes_\tau Y)$.

(iv) The unit principle holds if and only if $\Phi(\{\alpha\}) =_{\mathfrak{U}} 1_{\mathcal{I}} \iff \{i \in \mathcal{I} : |(\{\alpha\})_i| = 1\} \in \mathfrak{U} \iff \{i \in \mathcal{I} : 0 < |(\{\alpha\})_i|\} \in \mathfrak{U} \iff C_{\varnothing\{\alpha\}} \in \mathfrak{U}$

□

Remark again that the proof of the Cantor property is easy here, because of the adjusted property of the numerosity function, like mentioned before. For the weaker statement, proving necessity seems impossible, although in [2], equivalence is claimed.

In order to define a numerosity function, it suffices to define a finite approximation $\phi : \mathcal{W} \times \mathcal{I} \to \mathcal{W}$, and an ultrafilter $\mathfrak{U}$ over $\mathcal{I}$ that satisfies the properties of the above theorem.

Let now $\mathcal{I} = \mathscr{P}_{\text{fin}}(\mathit{Ord}) = \{i \in \mathscr{P}(\mathit{Ord}) : |i| \text{ finite }\}$, ordered by inclusion. Notice that $\langle \mathcal{I}, \subseteq \rangle$ is a directed class. Define a finite approximation as follows:

$$\phi(X, i) = X \cap i$$

The properties of definition 4.2.1 are obviously satisfied by this definition. So in order to provide a numerosity function on $\mathcal{W}$, what's left to show is that there exists an ultrafilter $\mathfrak{U}$ that contains the following family:

$$\mathcal{F} = \{C_{AB} : |A| < |B|\} \cup \{P^{\tau}_{AB} : \tau \text{ a } \theta\text{-tile, } A \subseteq \tau, B \subseteq \delta < \theta^{\omega}\}$$

So that all properties of a numerosity function are satisfied by theorem 4.2.4. From corollary A.1.4 follows that if $\mathcal{F} \subseteq \mathscr{P}(\mathcal{I})$ has the finite intersection property (recall definition 3.3.5), then there exists an ultrafilter on $\mathcal{I}$ that contains this family of sets as a subset. So if it is proven that the family $\mathcal{F}$ above has the finite intersection property, then this proves existence of a numerosity function.

This property of $\mathcal{F}$ is however not quite trivial. The proof in [2] it makes use of the following, very technical lemma.

Lemma 4.2.5. *Let $\theta_1 > \cdots > \theta_m = \omega$ be atoms, and $X_1, \ldots, X_n$ an arbitrary (finite) number of elements of $\mathcal{W}$. Then for every $k \in \mathbb{N}$, there exists $I \in \mathcal{I} = \mathscr{P}_{\mathit{fin}}(\mathit{Ord})$ such that the following holds:*

(i) Every finite $X_u \subset I$

(ii) For every infinite X_u:

$$|I \cap X_u| > |I \cap \cup\{X_v : |X_v| < |X_u|\}|$$

So an infinite X_u has more elements in common with I than the smaller X_v altogether.

(iii) For $s \leq m$, $h_1, \ldots, h_s \leq k$ integers, then if $X_u \subseteq \theta_1^{h_1} \ldots \theta_s^{h_s}$ and $X_v \subseteq \theta_s^k$, then:

$$(X_u \cap I) \otimes_{\theta_1^{h_1} \ldots \theta_s^{h_s}} (X_v \cap I) = (X_u \otimes_{\theta_1^{h_1} \ldots \theta_s^{h_s}} X_v) \cap I$$

The proof consists of a very technical construction of I. To improve readability, I've separated it using two more lemmas. First I introduce some notation. For α an infinite ordinal, let $\sum_i \omega^{\alpha_i} a_i$ be its normal form. Then an *atom of* α is any θ_{ij} appearing in any of the product expansion of the tiles $\omega^{\alpha_i} = \prod_j {\theta_{ij}}^{h_{ij}}$. The *degree* of α is defined as $\deg(\alpha) = \max_{i,j}\{a_i, h_{ij}\}$. For X a set of ordinals, define $\text{at}(X)$ as the set of all atoms of elements of X, and let $\text{span}(X) = \{\alpha : \text{at}(\{\alpha\}) \subseteq \text{at}(X)\}$. Finally, define $\text{span}_d(X) = \{\alpha \in \text{span}(X) : \deg(\alpha) < d\}$. The following lemma gives some useful properties of these new definitions.

Lemma 4.2.6. *Let X be a set of ordinals. If X is finite, then $at(X)$ is finite and $|span(X)| = \aleph_0$. If $|X| = \aleph_0$, then $|span(X)| = \aleph_0$ and $at(X) \leq \aleph_0$. $|X| = |at(X)| = |span(X)|$ whenever any of them is uncountable. Furthermore:*

$$|span_d(X)| = \begin{cases} d^{d^{|at(X)|}} & \text{if } at(X) \text{ is finite} \\ |at(X)| & \text{otherwise} \end{cases}$$

Proof. The first two statements just follow from combinatorics with finite or countably infinite many elements. Moreover, each of the three sets can only differ in size by a countable number of elements. So indeed $|X| = |\text{at}(X)| = |\text{span}(X)|$ whenever any of them is uncountable.

For $|\text{at}(X)|$ finite, consider an element $\alpha \in X$. For each ω^{α_i} in the normal form of α, a tile ω^{α_i} has $d^{|\text{at}(X)|}$ possible product expansions, As the degree can be at most $d-1$, and an atom can be absent (power 0). Each tile has d possible coefficients, from 0 to $d-1$. So α is one of $d^{d^{|\text{at}(X)|}}$ elements. The case when $|\text{at}(X)|$ is infinite follows in the same way, as each element of $\text{span}_d(X)$ is just a combination of finitely many elements of $\text{at}(X)$. ☐

Lemma 4.2.7. *Let $\theta_1 > \cdots > \theta_m = \omega$, $X_1, \ldots, X_u$ be as in lemma 4.2.5. Let $\kappa_1 < \cdots < \kappa_r$ an enumeration of the sizes of all uncountable X_u, $\kappa_0 = \aleph_0$, and define $\nu_i = \kappa_i$ if κ_i is regular, $\nu_i = {\kappa_{i-1}}^+$ otherwise. Define $Y_i = \bigcup\{X_u : |X_u| \leq \kappa_i\}, Z_i = \{X_u : |X_u| < \kappa_i\}$ and $\Theta = \{\theta_1, \ldots, \theta_m\}$. Let $A_i = \Theta \cup at(Y_i)$ and $B_i = \Theta \cup at(Z_i)$. Then there exists a finite $F_i \subset B_i$ and $d_i < \omega$ such that:*

$$|X_u \cap span_{d_i}(F_i \cup (A_i \setminus B_i))| \geq \nu_i \text{ for all } X_u \text{ with } |X_u| = \kappa_i$$

In the same way there is a finite $F_0 \subset A_0$ and $d_0 < \omega$ such that $Z_0 \cup \Theta \subseteq span_{d_0}(F_0)$ and

$$|X_u \cap span_{d_0}(F_0)| > |Z_0| \text{ for all } X_u \text{ with } |X_u| = \aleph_0$$

Proof. Let $i \in \{1, \ldots, r\}$ be arbitrary, and consider a set X_u with $|X_u| = \kappa_i$. Define for each finite $F \subseteq B_i$:

$$X_u^F = \{\alpha \in X_u | \text{at}(\{\alpha\}) \cap B_i \subseteq F\}$$

Then:

$$X_u = \bigcup_{F \in \mathscr{P}_{\text{fin}}(B_i)} X_u^F$$

Notice $|B_i| = |\kappa_{i-1}|$ (follows from lemma 4.2.6). Furthermore $\mathscr{P}_{\text{fin}}(B_i)$ is the union over all n of subsets of length n. Each such subset has cardinality $\kappa_{i-1} \geq \aleph_0$ ($(\kappa_{i-1})^n = \kappa_{i-1}$ by lemma 2.2.5), hence $|\mathscr{P}_{\text{fin}}(B_i)| = \omega \cdot \kappa_{i-1} = \kappa_{i-1}$.

But $|X_u| = \kappa_i \geq \nu_i > \kappa_{i-1}$, with ν_i regular (any successor is regular), so it is not the union of smaller cardinals. By the above and lemma 2.2.7, it follows that there must be a finite F_u such that $|X_u^{F_u}| \geq \nu_i$. As $X_u^{F_u} = \bigcup_{d<\omega}\{\alpha \in X_u^{F_u} | \deg(\alpha) \leq d\}$, it holds with the same argument that there exists a $d_u < \omega$ such that:

$$|X_u^{F_u} \cap \text{span}_{d_u}(X_u^{F_u})| \geq \nu_i$$

By definition $X_u^{F_u}$ is the set of ordinals of which only finitely many atoms are in B_i. The rest must therefore be in A_i. So the above can be restricted further:

$$|X_u^{F_u} \cap \text{span}_{d_u}(A_i \setminus B_i)| \geq \nu_i$$

Now let $F_i = \bigcup\{F_u : |X_u| = \kappa_i\}, d_i = \max\{d_u : |X_u| = \kappa_i\}$. Using $X_u^{F_i} \cap \operatorname{span}_{d_i}(A_i \smallsetminus B_i) \subseteq X_u \cap \{\operatorname{span}_{d_i}(F_i \cup (A_i \smallsetminus B_i))\}$, we end up with:

$$|X_u \cap \{\operatorname{span}_{d_i}(F_i \cup (A_i \smallsetminus B_i))\}| \geq \nu_i$$

Where F_i is finite, d_i an integer, so this proves the first assertion.

For the second assertion, let X_u with $|X_u| = \aleph_0$ be arbitrary. We proceed by first defining a finite F_u and an integer d_u so that the following holds for this particular X_u:

$$|X_u \cap \operatorname{span}_{d_u}(F_u)| > |Z_0| \text{ with } Z_0 \cup \Theta \subseteq \operatorname{span}_{d_u}(F_u)$$

Let $d_m = \max\{\deg(\alpha) : \alpha \in Z_0\}$ (Z_0 is finite). Also B_0 is finite, so for $Z_0 \cup \Theta \subseteq \operatorname{span}_{d_u}(F_u)$ to holds we can just require $B_0 \subseteq F_u$, $d_u \geq d_m$. Now either X_u contains elements of degree $> d_m$, or $A_0 \smallsetminus B_0 \neq \varnothing$ (or both). As $|Z_0|$ is finite, it is therefore a matter of enlarging d_u or extending F_u with elements from $A_0 \smallsetminus B_0$ to satisfy the above equation.

With $F_0 = \bigcup\{F_u : |X_u| = \aleph_0\}$ and $d_0 = \max\{d_u : |X_u| = \aleph_0\}$ the assertion is proved. ☐

The proof of lemma 4.2.5 is now as follows:

Proof. Let $Y_i, Z_i, A_i, B_i, \Theta, \nu_i, d_i, F_i$ be as in lemma 4.2.7 (for $i \in \{1, \dots, r\}$), and $k \in \mathbb{N}$ arbitrary. Let $d = \max\{2k, d_0, \dots, d_r\}$ and $F = \bigcup_{i=0}^r F_i$. The claim is that one can choose recursively on $i \in \{1, \dots, r\}$ finite subsets $G_i \subset A_i \smallsetminus B_i$ such that:

$$|X_u \cap \operatorname{span}_{d_i}(F_i \cup G_i)| > d^{d^{|H_{i-1}|}}$$

Where $H_s = F \cup \bigcup_{j=1}^s G_j$, $H_0 = F$. Notice $d^{d^{|H_{i-1}|}} = |\operatorname{span}_d(H_{i-1})|$, with $H_{i-1} \subset F \cup B_i$. From lemma 4.2.7 we know $|X_u \cap \{\operatorname{span}_{d_i}(F_i \cup (A_i \smallsetminus B_i))\}| \geq \nu_i$. Also F contains only finitely many elements. It follows that it must be possible to extend F_i with a finite set G_i containing new elements from $A_i \smallsetminus B_i$ such that the above set reaches the desired size. Because for each i the set $|X_u \cap \{\operatorname{span}_{d_i}(F_i \cup (A_i \smallsetminus B_i))\}|$ is uncountably bigger (by a regular cardinal) then the previous set (for $i-1$), this is certainly possible.

Let $I = \operatorname{span}_d(H_r)$. The only thing left is to verify the properties (i)-(iii).

Property (i) holds as $\bigcup\{X_u : |X_u| \text{ finite }\} \subseteq Z_0 \cup \Theta \subseteq \operatorname{span}_{d_0}(F_0) \subseteq I$.

For $|X_u| = \kappa_i$, property (ii) follows from: $\{X_v : |X_v| < |X_u|\} \cap I = Z_i \cap I \subseteq \operatorname{span}(B_i) \cap I \subseteq \operatorname{span}_d(H_{i-1})$, thus $|\{X_v : |X_v| < |X_u|\}| \leq |\operatorname{span}_d(H_{i-1})| = d^{d^{|H_{i-1}|}} < |X_u \cap \operatorname{span}_{d_i}(F_i \cup G_i)| \leq |X_u \cap I|$. For $|X_u| = \kappa_0 = \aleph_0$, it follows from $|\{X_v : |X_v| < |X_u|\} \cap I| = |Z_0 \cap I| < |(X_u \cap \operatorname{span}_{d_0}(F_0)) \cap I| \leq |X_u \cap I|$, by the second assertion of lemma 4.2.7.

At last, let $s \leq m, h_t \leq k$ for $1 \leq t \leq s$, $\tau = \theta_1^{h_1} \dots \theta_s^{h_s}$. For any $A \supseteq \Theta$, notice that, because $d \geq 2k$ and the degree of any element of $\tau \cap \operatorname{span}_d(A) \otimes_\tau \theta_s^k \cap \operatorname{span}_d(A)$ cannot exceed $2k$:

$$(\tau \cap \operatorname{span}_d(A)) \otimes_\tau (\theta_s^k \cap \operatorname{span}_d(A)) \subseteq \operatorname{span}_d(A)$$

And notice there's no absorption in the product, for $\theta_s^k \subseteq \theta_s^\omega$ (τ is a θ_s-tile). So indeed, for $X_u \subseteq \tau, X_v \subseteq \theta_s^k$ and using $\Theta \in H_r$:

$$(X_u \cap I) \otimes_\tau (X_v \cap I) = (X_u \otimes_\tau X_v) \cap I$$

Which completes the proof of the lemma. ☐

The lemmas above have done all the technical work to prove the following theorem, that has already been announced.

Theorem 4.2.8. *The set* $\mathcal{F} = \{C_{AB} : |A| < |B|\} \cup \{P^{\tau}_{AB} : \tau$ *a* θ*-tile,* $A \subseteq \tau, B \subseteq \delta < \theta^{\omega}\}$ *has the finite intersection property. Therefore, there exists an ultrafilter* $\mathfrak{U}$ *on* $\mathcal{I}$ *that justifies the existence of a numerosity function on* $\mathcal{W}$.

Proof. It has already been mentioned that the finite intersection property is the only thing to prove. So, let $F_1, \dots, F_l$ be a finite family of sets $\in \mathcal{F}$. These sets are all of the form $C_{A_tB_t}$ or $P^{\tau_t}_{A_tB_t}$ for $1 \leq \tau \leq l$. Let $X_1, \dots, X_n$ be an enumeration of the sets A_t, B_t appearing in $F_1, \dots, F_l$. We'd like to apply lemma 4.2.5 to these sets. We set up k, $\theta_1 > \theta_2 > \dots > \theta_m = \omega$ now in such a way that whenever $F_t = P^{\tau_t}_{A_tB_t}$ for some t, then:

$$\begin{aligned} \tau_t &= \theta_1^{h_{1t}} \dots \theta_m^{h_{mt}} \text{ where } 0 \leq h_{st} \leq k \text{ for all } s \leq m \\ B_t &\subseteq \theta_s^k \text{ where s is the largest index with } h_{st} > 0 \end{aligned}$$

So in fact $\theta_1, \dots, \theta_m$ are just all atoms that occur in all τ's of this finite family of sets.

With this, lemma 4.2.5 is applicable. Consider first an element of the form $C_{A_tB_t}$. Then $|A_t| < |B_t|$, so in case B_t is infinite, part (ii) of the lemma gives $|A_t \cap I| < |B_t \cap I|$ with $I \in \mathcal{I}$, hence $I \in C_{A_tB_t}$ by definition. If B_t is finite, then $|A_t| < |B_t|$ implies $|A_t \cap I| < |B_t \cap I|$, as $X_u \cap I = X_u$ for every finite X_u. So again $I \in C_{A_tB_t}$.

For an element of the form $P^{\tau_t}_{A_tB_t}$, the variables above are chosen in such a way that the lemma gives:

$$A_t \cap I \otimes_{\tau_t} B_t \cap I = (A_t \otimes_{\tau_t} B_t) \cap I$$

There is no absorption (A_t and B_t satisfy the appropriate conditions), and I is finite, so:

$$|A_t \cap I \otimes_{\tau_t} B_t \cap I| = |A_t \cap I| \cdot |B_t \cap I| = |(A_t \otimes_{\tau_t} B_t) \cap I|$$

Or $|(A_t)_I| \cdot |(B_t)_I| = |\left((A_t \otimes_{\tau_t} B_t)\right)_I|$. So by definition, $I \in P^{\tau_t}_{A_tB_t}$.

Thus $I \in \bigcap_{t \leq l} F_l$, hence $\mathcal{F}$ satisfies the finite intersection property. □

At this point, we have proven that an ultrafilter $\mathfrak{U}$ that satisfies the necessary conditions from theorem 4.2.4 to provide a numerosity function $\mathfrak{n}$ on $\mathcal{W}$ exists. In [2], this is summarized in the following corollary:

Corollary 4.2.9. *There exist numerosity functions* $\mathfrak{n} : \mathscr{P}(Ord) \to \mathcal{A}$ *where* $\mathcal{A}$ *is a class of hyperintegers in the sense of nonstandard analysis.*

Proof. The existence has been proven, the qualification of $\mathcal{A}$ is a bit dangerous. I have remarked that $\mathcal{A} = \mathbb{Z}^{\mathcal{I}}/\mathfrak{U}$ can be seen as an ultrapower of $\mathbb{Z}$. When $\mathbb{Z}^{\mathcal{I}}$ can be considered a set-ring (again, I refer to appendix A.2 for this case), then by the ultrapower construction of a nonstandard embedding (like mentioned in section 3.4) it can indeed be considered a "class" of hyperintegers. □

4.3 Back to the countable case

One might ask at this point, what is the link between the numerosity defined here, and the numerosity defined in chapter 3. To answer that question, we can restrict the numerosity function to set of all countable sets of ordinals: $\mathcal{W} = \mathscr{P}(\omega)$ in that case. Also the class $\mathcal{I}$ can be restricted: define $\mathcal{I} = \mathscr{P}_{\text{fin}}(\omega)$. So $\mathfrak{U}$ is an ultrafilter on the restricted $\mathcal{I}$. Notice the "labelling" is the same as the canonical labelling of $\mathbb{N}$, so at this point it is possible to compare the numerosity function that is defined on subsets of $\mathbb{N}$ with the general numerosity function defined here.

In chapter 3, a set $A \subset \mathbb{N}$ is compared to a set $B \subset \mathbb{N}$ by using finite approximations and an ultrafilter on $\mathbb{N}$. Denote that ultrafilter (to distinguish it from the ultrafilter on $\mathcal{I}$) by $\mathfrak{U}_{\mathbb{N}}$.

In the notation of this chapter, let $I_i = \{1, 2, \dots, i\}$, then for the labelled sets $\mathfrak{n}(\mathbf{A}) = \mathfrak{n}(\mathbf{B}) \iff \{i : A \cap I_i = B \cap I_i\} \in \mathfrak{U}_{\mathbb{N}}$. So the two notions of size coincide when for $\mathscr{I} = \{I_1, I_2, \dots\}$ holds:

$$\begin{aligned} &\{I \in \mathfrak{U} : A \cap I = B \cap I\} \cap \mathscr{I} \in \mathfrak{U} \iff \{i : A \cap I_i = B \cap I_i\} \in \mathfrak{U}_{\mathbb{N}} \text{ or} \\ &\{I_i \in \mathfrak{U} : A \cap I_i = B \cap I_i\} \in \mathfrak{U} \iff \{i : A \cap I_i = B \cap I_i\} \in \mathfrak{U}_{\mathbb{N}} \end{aligned}$$

So in general, the notions of size coincide for all subsets of $\mathbb{N}$ when for any set $J \subset \mathbb{N}$ holds:

$$\{I_j : j \in J\} \in \mathfrak{U} \iff J \in \mathfrak{U}_{\mathbb{N}}$$

Whether it is possible to establish ultrafilters $\mathfrak{U}$ on $\mathscr{P}_{\text{fin}}(\omega)$ and $\mathfrak{U}_{\mathbb{N}}$ on $\mathbb{N}$ that satisfy this relation, and still satisfy all other properties that were required, I cannot figure out yet.

It is not needed though, as there is a theory for the countable case. But maybe some thoughts on this might enlarge the set of sets on which the numerosity function of chapter 3 can be defined. For example, the difference principle is not a requirement in the numerosity function in this chapter:

$$\xi < \mathfrak{n}(\mathbf{A}) \iff \exists \mathbf{B} \subset \mathbf{A}, \mathfrak{n}(\mathbf{B}) = \xi$$

So it is certainly interesting to investigate. In [1] and [2] it is left as an open question whether this difference principle might be extended, and here I will do the same, from a different point of view.

Appendix A

Selected topics

A.1 Ultrafilters

This section provides a short introduction to the concept of filters, and states a few results that are used in the chapters above.

Definition A.1.1. *A **filter** on a set* $S \neq \varnothing$ *is a set* $\mathfrak{F} \subset \mathscr{P}(S)$ *such that:*

(i) $S \in \mathfrak{F}$ *and* $\varnothing \notin \mathfrak{F}$.

(ii) If $X \in \mathfrak{F}$ *and* $Y \in \mathfrak{F}$*, then* $X \cap Y \in \mathfrak{F}$.

(iii) If $X \in \mathfrak{F}$ *and* $Y \supset X$*, then* $Y \in \mathfrak{F}$.

A filter $\mathfrak{U}$ *is called an **ultrafilter** if for any* $X \subset S$ *holds* $X \in S$ *or* $S \setminus X \in S$.

The following lemma is used in theorem 4.2.8:

Lemma A.1.2. *Let* S *be a set. If a family of sets* $\mathcal{F} \subset \mathscr{P}(S)$ *has the finite intersection property, then there exists a filter* $\mathfrak{F}$ *on* S *such that* $\mathcal{F} \subset \mathfrak{F}$.

Proof. Let $\mathfrak{F}$ defined by all sets $D \in \mathscr{P}(S)$ such that there is a finite subset $\{X_1, \ldots, X_n\} \in \mathcal{F}$ with $X_1 \cap \cdots \cap X_n \subseteq D$. Then property (i) of definition A.1.1 holds, as $\mathcal{F}$ has the finite intersection property.

If $D, E \in \mathfrak{F}$, pick the finite $\{X_1, \ldots X_n\}$ and $\{Y_1, \ldots, Y_m\}$ such that $X_1 \cap \cdots \cap X_n \subseteq D$ and $Y_1 \cap \cdots \cap Y_m \subseteq E$. Then $X_1 \cap \cdots \cap X_n \cap Y_1 \cap \cdots \cap Y_m \subseteq D \cap E$ so $D \cap E \in \mathfrak{F}$ by definition.

Property (iii) is trivial from the definition.

So $\mathfrak{F}$ is a filter, and $\mathcal{F} \subset \mathfrak{F}$. ☐

In theorem 4.2.8 an ultrafilter was needed, however. There is needed a bit more on ultrafilters. A filter $\mathfrak{F}$ on a set S is called *maximal* if there is no filter $\mathfrak{F}'$ on S with $\mathfrak{F}' \supset \mathfrak{F}$. It is easily seen that any ultrafilter $\mathfrak{U}$ is maximal: if there is $D \in \mathfrak{U}' \setminus \mathfrak{U}$, then $S \setminus D$ was already in $\mathfrak{U}$, so $\varnothing = D \cap (S \setminus D) \in \mathfrak{U}'$, a contradiction.

The converse also holds, suppose a filter $\mathfrak{F}$ on S is not an ultrafilter, so there is a $D \subset S$ with both D and $S \setminus D$ not in $\mathfrak{F}$. Then for any $E \in \mathfrak{F}$, $E \cap D \neq \varnothing$, for otherwise $E \subset S \setminus D$, so then $S \setminus D \in \mathfrak{F}$. It follows that the set $\mathcal{G} = \mathfrak{F} \cup \{D\}$ has the finite intersection property. By lemma A.1.2, there is a filter $\mathfrak{F}' \supseteq \mathcal{G} \supset \mathfrak{F}$.

So an ultrafilter is the same thing as a maximal filter. Therefore, the following lemma can be proven using Zorn's lemma:

Lemma A.1.3. *Any filter $\mathfrak{F}$ on a set S can be extended to an ultrafilter $\mathfrak{U}\sup\mathfrak{F}$.*

Proof. Define the set $P = \{F \subset \mathscr{P}(S) : F \text{ a filter with } F \supseteq \mathfrak{F}\}$, partially ordered by inclusion $\subseteq$. For C a chain in P, notice that $\bigcup C$ is a filter on S, and $\bigcup C$ is an upper bound for any filter in C. So any chain in P has an upper bound, then by Zorn's lemma P has a maximal element $\mathfrak{U}$. This is then an ultrafilter on S containing $\mathfrak{F}$. ☐

For completeness, I state the following corollary, which is precisely what is used in theorem 4.2.8.

Corollary A.1.4. *If a family of sets $\mathcal{F} \subset \mathscr{P}(S)$ has the finite intersection property, then there exists an ultrafilter $\mathfrak{U}$ on S such that $\mathcal{F} \subset \mathfrak{U}$.*

A.1.1 Ultraproducts

An important application of ultrafilters is giving a rather easy proof of the existence of a nonstandard embedding $\langle \mathscr{V}(X_0), \mathscr{V}(Y_0), *\rangle$, and even giving an elegant proof of the upwards Löwenheim-Skolem theorem. What is used in both cases is an ultraproduct of certain models M_i.

For A_i a family of sets indexed by I, and $\mathfrak{U}$ an ultrafilter on I, the *ultraproduct* of the A_i with respect to $\mathfrak{U}$ is defined as the equivalence classes of all sequences ϕ with $\phi_i \in A_i$ under $\mathfrak{U}$. That is, $\phi =_{\mathfrak{U}} \psi \iff \{i : \psi_i = \phi_i\} \in \mathfrak{U}$. The ultraproduct is denoted as follows:

$$\prod_{\mathfrak{U}} A_i = \{[\phi]_{\mathfrak{U}} : \phi \text{ a sequence with } \phi_i \in A_i \text{ for each } i \in I\}$$

From a family of models M_i of a certain language L, a model $\mathcal{M}$ can be constructed as $\mathcal{M} = \prod_{\mathfrak{U}} M_i$. Relations, functions and constants are defined "modulo" the ultrafilter:

(i) For a relation symbol R:

$$R^{\mathcal{M}}(\phi_1, \dots, \phi_n) \iff \{i : R^{M_i}(\phi_1(i), \dots, \phi_n(i))\} \in \mathfrak{U}.$$

(ii) For a function symbol F,

$$F^{\mathcal{M}}(\phi_1, \dots, \phi_n) = [i \mapsto F^{M_i}(\phi_1(i), \dots, \phi_n(i))]_{\mathfrak{U}}$$

(iii) For a constant symbol c:

$$c^{\mathcal{M}} = [i \mapsto c^{M_i}]_{\mathfrak{U}}$$

Of course it should be checked that this is well-defined (it should not depend on the choice of representatives $\phi_1, \dots, \phi_n$ for example). It is the case, but I don't go into much detail here.

The following theorem can be proven in a way similar to the proof of theorem 3.4.7.

Theorem A.1.5 (The Fundamental theorem of Ultraproducts)**.** *Let $\mathcal{M} = \prod_{\mathfrak{U}} M_i$ be an ultraproduct of L−models M_i. Then:*

(i) For any term $t(x_1,\ldots,x_n)$ and $\phi_1,\ldots,\phi_n \in \mathcal{M}$, we have:

$$t^{\mathcal{M}}(\phi_1,\ldots,\phi_n) = [i \mapsto t^{M_i}(\phi_1(i),\ldots,\phi_n(i))]_{\mathfrak{U}}$$

(ii) For any L–formula $\sigma(x_1,\ldots,x_n)$ and $\phi_1,\ldots,\phi_n \in \mathcal{M}$, we have:

$$\mathcal{M} \models \sigma(\phi_1,\ldots,\phi_n) \iff \{i : M_i \models \sigma(\phi_1(i),\ldots,\phi_n(i))\} \in \mathfrak{U}$$

(iii) For any L–sentence σ:

$$\mathcal{M} \models \sigma \iff \{i : M_i \models \sigma\} \in \mathfrak{U}$$

The proof can be found in [3].

If $M_i = M$ for all i, then $\prod_{\mathfrak{U}} M$ is called the *ultrapower* of M with respect to $\mathfrak{U}$. On page 23 it is shown that the set of numerosities $\mathscr{N}$ is an ultrapower of $\mathbb{N}$.

A.2 Foundation of proper classes within ZFC

In this short section, I will describe a kind of workaround that justifies the existence of proper classes and the structures I assumed in defining for example $\mathbb{Z}^{\mathcal{I}}$, where $\mathcal{I}$ is a proper class. Most of this section is adopted from [4].

In the metalanguage of ZFC (so talking about statements in ZFC), a class basically represents a formula. For $\psi(x)$ a formula that says: x has property ψ, one can think of $\{x : \psi(x)\}$ as the class of all (sets) x with property ψ, without quantifying over a set. The class $\{x : x \text{ is an ordinal }\}$ is a proper class defined this way. In this framework, it is not clear what $\mathbb{Z}^{\mathcal{I}}$ is supposed to be.

It is certainly possible to go beyond ZFC to solve this problem to a certain extend, there is however a more or less satisfying solution that can be formulated within ZFC. To state it, a notion of set theory called the *cumulative hierarchy of sets* is needed.

This notion involves the axiom of regularity, or foundation of ZFC. It says that any nonempty set S has a $\in$-minimal element:

$$\forall S : S \neq \varnothing \rightarrow \exists (x \in S) : S \cap x = \varnothing$$

It is an axiom that does not interfere with the other axioms. Basically all mathematics can be founded in ZFC without this axiom. It solves some issues though, for instance that a set cannot contain itself, there are no chains $x_0 \in x_1 \in \cdots \in x_0$, and more of this kind. More importantly, it is used for constructing models of set theory, and talking about ZFC in a metalanguage, and that is precisely the goal here.

For S a set, define the *transitive closure* of S as $TC(S) = \bigcap\{T \supset S : T \text{ transitive }\}$. Note that for any transitive set T, $\bigcup T \subseteq T$. So if one lets $S_0 = S$, $S_{n+1} = \bigcup S_n$, it follows $TC(S) = \bigcup_n S_n$ (this gives also the existence of $TC(S)$). The following lemma might already hint to the meta-languistic implications of the axiom of foundation.

Lemma A.2.1. *Any class of sets C has a $\in$-minimal element.*

Proof. Pick $X \in C$ arbitrary. If $X \cap C = \varnothing$, then we're done. Otherwise, let $T = TC(X \cap C)$ (note $X \cap C$ is a set). Then $Y = T \cap C$ has a $\in$-minimal element y by the axiom of foundation, so $y \cap Y = \varnothing$. Then $y \cap C = \varnothing$, for otherwise, any $z \in y \cap C$ implies $z \in T$ (T is transitive), so $z \in T \cap C$ which is a contradiction. $\square$

The *cumulative hierarchy of sets* is an indexed family V_α, where α is an ordinal, defined as follows:

$$\begin{array}{l} V_0 = \varnothing; V_{\alpha+1} = \mathscr{P}(V_\alpha) \\ V_\alpha = \bigcup_{\beta<\alpha} V_\beta \text{ if } \alpha \text{ is a limit ordinal} \end{array}$$

Notice that each V_α is transitive, $\alpha < \beta$ implies $V_\alpha \subset V_\beta$, and $\alpha \subset V_\alpha$. The following proposition shows even more of the remarkable properties of the axiom of foundation.

Proposition A.2.2. *For every set x there is an α such that $x \in V_\alpha$.*

Proof. Let C be the class of sets for which such α does not exist, and assume $C \neq \varnothing$. Then there is a least element x of C, by lemma A.2.1. So, for any $y \in x$, $y \in V_{\alpha_y}$ for some α_y. Therefore $\{\alpha_y : y \in x\}$ is a set of ordinals. Let $\gamma = \sup\{\alpha_y : y \in x\}$. Then for any y, $V_{\alpha_y} \subset V_\gamma$. Hence $x \subset V_\gamma$, which implies $x \in V_{\gamma+1}$, a contradiction. So $C = \varnothing$, which completes the proof. $\square$

This proposition shows that V_α is indeed a cumulative hierarchy of sets. In fact $\bigcup_{\alpha \in Ord} V_\alpha$ can be considered the class of all sets. For each set x, the rank of x is the least ordinal α such that $x \subset V_\alpha$ (so $x \in V_{\alpha+1}$).

An interesting result is obtained by assuming a strongly inaccessible cardinal, which will be defined first.

Definition A.2.3. *A cardinal κ is called **strongly inaccessible** if:*

(i) κ is uncountable

(ii) κ is regular (see page 12)

(iii) κ is a strong limit, that is, if $\mu < \kappa$, then $2^\mu < \kappa$.

Notice $\aleph_0$ only fails for property (i). Strongly inaccessible cardinals cannot be obtained from smaller cardinals by ordinary set theoretic operations such as taking powerset, multiplication and addition. It cannot be shown that the existence of strongly inaccessible cardinals is consistent with ZFC [1]. However, there is the following result:

Proposition A.2.4. *For κ a strongly inaccessible cardinal, V_κ is a model of ZFC. Furthermore:*

$$\begin{array}{l} \alpha \text{ is an ordinal} \iff \alpha \text{ is an ordinal in } V_\kappa \\ \alpha \text{ is a (regular) cardinal} \iff \alpha \text{ is a (regular) cardinal in } V_\kappa \end{array}$$

Stating a proof would go beyond the scope of this text, but it can be found in [4] as well.

So what can be done to solve the issue stated in the beginning of this section is assuming a "meta-model" of ZFC (consisting of the cumulative hierarchy),

---

[1] See [4], *"Inaccessibility of Inaccessible Cardinals"* in chapter 12

in which a strongly inaccessible cardinal κ exists. That provides us with a "smaller" model V_κ, where all set theory that is done here (on ordinals, cardinals and numerosities) can be done. However, in the meta-model, the defined classes can be treated as sets, so the class of labelled sets $\mathscr{L}$, or the class of ordinals *Ord* are just subsets of V_κ. Also $\mathbb{Z}^{\mathcal{I}}$ can be defined within ZFC in that way, while all notions of ordinals and cardinals correspond in the two models.

In short, the assumption of a strongly inaccessible cardinal is a proper foundation for all classes and class-rings that are defined here. The reader might find this a dissatisfying solution. However, belief in the existence of strongly inaccessible cardinals is widely shared among present-day set theorists. There are many theorems in mathematics that rely on such assumptions.

Bibliography

[1] Vieri Benci and Mauro Di Nasso. Numerosities of labelled sets: a new way of counting. *Adv. Math.*, 173(1):50–67, 2003.

[2] Vieri Benci, Mauro Di Nasso, and Marco Forti. An Aristotelian notion of size. *Ann. Pure Appl. Logic*, 143(1-3):43–53, 2006.

[3] C. C. Chang and H. J. Keisler. *Model theory*, volume 73 of *Studies in Logic and the Foundations of Mathematics.* North-Holland Publishing Co., Amsterdam, third edition, 1990.

[4] Thomas Jech. *Set theory.* Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2003. The third millennium edition, revised and expanded.

[5] Kenneth Kunen. *Set theory*, volume 102 of *Studies in Logic and the Foundations of Mathematics.* North-Holland Publishing Co., Amsterdam, 1980. An introduction to independence proofs.

[6] Saharon Shelah. *Proper and improper forcing.* Perspectives in Mathematical Logic. Springer-Verlag, Berlin, second edition, 1998.

[7] K. D. Stroyan and W. A. J. Luxemburg. *Introduction to the theory of infinitesimals.* Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1976. Pure and Applied Mathematics, No. 72.