An Introduction to Nonstandard Analysis

14th June 2020

A TWIN Bachelor Thesis (15EC)

Shane de Jong

Utrecht University

Supervisor:

Dr. Jaap van Oosten
Mathematical Institute, Utrecht University

Infinitesimal Beginnings

Where at the foundation of contemporary mathematics lies logic and set theory, in the 16th century is found geometry. As relic of the ancient Greeks, geometry is one of the few fields that survived the scrutiny of the 16th century scientist. Whereas Aristotle's theory of gravity, or Ptolemy's cosmological framework were no longer deemed true, the results of Euclidean geometry seemed to have survived the nearly 2000 year journey into the Renaissance. The 16th century scientist considered this, and concluded that there was some fundamental truth in the way Euclidean geometry is done. Geometry therefore found itself placed at the base of scientific, and mathematical reasoning.

Passing into the 17th century, scientists had to reconcile the fact that their questions had changed but their geometric reasoning had not. Where the Greeks dealt with squares and circles, now there were transcendental curves to consider. It is here we find Newton and Leibniz, founders of calculus, and their usage of *infinitesimals*. As an element both infinitely small, but not actually zero, it offered its user the ability to use stick with geometric arguments, and apply them on the new types shapes and curves that came with the 17th century.

This geometric approach, together with the use of infinitesimals, continued into the 19th century. Nearing the second half of this century however, problems for the Greek tradition start to emerge. Geometry had found its place through the inherent truth it represented, but new developments, especially those in non-Euclidean geometry, raised very valid concerns as to its central philosophical position. For infinitesimals similar concerns were raised. These were not new, as even its earliest usage had seen its methods questioned. As its most famous critic, the 17th century bishop George Berkeley, said:

> *"And what are these same evanescent increments? They are neither finite quantities nor quantities infinitely small, nor yet nothing. May we not call them the ghosts of departed quantities?"*[1]

Infinitesimals, while intuitive in nature, lacked any sort of formalism. And questions raised such as those by Berkeley, were left unanswered.

A gradual shift started; new methods with stronger foundation, offered by Weierstrass and Cauchy, were able to entirely replace the need for infinitesimals. The second half of the 19th century saw the rise of the mathematicians need for formalism, and with it the decline of geometry and the use of infinitesimals.

This was not the end for infinitesimals however. In the same period of time we see the division of science in the fields of today. Where before the 19th century there was the scientist, afterwards there are physicists, mathematicians, biologists and chemists. So while mathematicians turned to formalism and its foundations – and with it disposed of infinitesimals – other fields, such as physics, did not.

It is therefore not uncommon, to come across infinitesimal arguments in the physics textbooks of today. While this might disturb the modern mathematician, the physicist possess a powerful technique: they have the ability to verify their answers with reality. And if some methods produces results, regardless of its mathematical rigour, it is a useful one.

Mathematicians do not possess such luxury, but even for them, the remarkable ability for infinitesimals to produce exact results is an intriguing one. Luckily for them there is *nonstandard analysis*. Originated in the early 1960s by the mathematician Abraham Robinson, its methods offer modern mathematical rigour for the concepts of infinitesimals.

The Present Work

In this text I offer an introduction to the subject of nonstandard analysis. Texts on the topic are usually divided in the constructionist approach, and the axiomatic approach. The present work leans heavily towards the constructionist side.

In the first chapter we introduce the concept of nonstandard reals. It is here we will find the infinitesimal elements. As is so often the case in mathematics however, it turns out that the methods of the first chapter can be generalised. The second and third chapter deal with this generalisation. In the second chapter we find the general construction of the nonstandard version of any set, not just the reals, while in the third we identify what the 'nice' sets of our new construction are.

The first chapter is therefore very much an introductory one, whose structure is later mirrored by chapters two and three.

[1] Cf. [1, p. 18].

Disclaimer

This text was written such that it would conclude with the concept known as the *Transfer Principle*. Because of time constraints I was not able to include a satisfactory discussion on the topic, and it is therefore not found in this text. In my opinion the transfer principle is a defining characteristic of nonstandard analysis, and since I was not able to discuss it, this text is lesser for it. As such the text ends rather abruptly after defining what are known as *internal elements*. The transition between first and second chapter is also a rough one, as I had planned to include more of the properties of the nonstandard reals.

Still I think the present text has merit. It introduces a lot of the concepts and techniques found in nonstandard analysis and does set up for the transfer principle. Therefore, using the present text as an introductory one and then finding a discussion on the transfer principle should be a smooth transition.

Finally, while not explicitly cited anywhere in the rest of the text, the article *An Invitation to Nonstandard Analysis* by Tom Lindstrøm formed both a crucial part in my own understanding of nonstandard analysis, as well as served as a basis for this text.[2]

[2] Cf. [3].

Contents

Notation and Terminology

- Sequences:
  - $\{a_n\}$ is shorthand for $(a_n)_{n\geq 0}$.
- Logic:
  - 'or' is inclusive.
- Relations and operations of sets:
  - (Inclusion) $A \subset B$ means that A is contained in B but not equal.
  - (Complement) $A - B$ is the set of all elements in A which are not in B.
  - (Disjoint Union) If $A \cap B = \emptyset$, then we will sometimes write $A \mathbin{\dot\cup} B$ for the union $A \cup B$ to stress that A and B are disjoint.
  - In general we will adopt the set $\to$ collection $\to$ family hierarchy. So a collection is a set of sets, and a family is a set of collections. This is to avoid sentences such as "A is a set of sets of sets".
  - In general we will say A is contained in B if $A \subseteq B$. If $A \in B$ we will say that A is an element of B instead.

The Nonstandard Reals

1.1 Construction

To construct the set of non-standard reals ${}^*\mathbb{R}$, we will first look at a similar and more familiar construction: the construction of the set of real numbers through the use of rational Cauchy-sequences. This construction starts by defining an equivalence relation on the set of all rational Cauchy-sequences, and defines addition, multiplication and a total order on the resulting set of equivalence classes componentwise; an approach we will see mirrored in the construction of ${}^*\mathbb{R}$.

To formalize this construction: let $S_{\mathbb{Q}}$ be the set of all rational Cauchy-sequences and $\equiv$ be an equivalence relation on $S_{\mathbb{Q}}$, such that

$$\{a_n\} \equiv \{b_n\}\,,\ \text{precisely when}\ \lim_{n\to\infty}(a_n - b_n) = 0.$$

Then the real numbers can be identified as the set of equivalence classes, $S_{\mathbb{Q}}/\equiv$. If we let $[a_n]$ denote the equivalence class of $\{a_n\}$ we can also define addition and multiplication as

$$[a_n] + [b_n] = [a_n + b_n],\ \text{and}\ [a_n]\cdot[b_n] = [a_n \cdot b_n].$$

Additionally, we can define a strict total order on $S_{\mathbb{Q}}/\equiv$ by letting $[a_n] < [b_n]$ when there is a positive rational number ϵ such that $\epsilon < b_n - a_n$ for all sufficiently large n. In other words: one element is smaller than another, when the tails of the representatives share the same relation componentwise.

Finally, there is a natural way of embedding the rational numbers in this identification of the real numbers through the function $q \mapsto [q]$ for all $q \in \mathbb{Q}$.

Naturally, someone unfamiliar with this construction of $\mathbb{R}$, should check whether $\equiv$ is actually an equivalence relation, and addition, multiplication and the total order are well defined. Having verified these facts, we can show that this interpretation of the real numbers is, in fact, an ordered field by transferring these properties from the rational numbers. Additionally we can show that $S_{\mathbb{Q}}/\equiv$ is Cauchy-complete by explicitly showing that the limit of a Cauchy-sequence in $S_{\mathbb{Q}}/\equiv$ exists. This involves creating a rational Cauchy-sequence first, and showing that this is the required limit.[1] Therefore this construction leads to a Cauchy-complete ordered field $S_{\mathbb{Q}}/\equiv$, that extends our original set of the rational numbers $\mathbb{Q}$. Finally, we can show that the Archimedean property of the rational numbers transfers to $S_{\mathbb{Q}}/\equiv$ as well. So $S_{\mathbb{Q}}/\equiv$ is an Archimedean Cauchy-complete ordered field.

An alternative characterisation of 'Archimedean and Cauchy-complete' is the equivalent concept of Dedekind-completeness. The latter states that any subset that is bounded from above, has a least upper bound. This equivalence is proven in Section A.1 of the appendix. Furthermore, any Dedekind-Complete ordered field is isomorphic to the real numbers as can be seen in Theorem A.1.14. Since $S_{\mathbb{Q}}/\equiv$ is a Dedekind-complete ordered field, we have constructed $\mathbb{R}$ from $\mathbb{Q}$ up to interpretation (isomorphism).

The construction of the ordered field ${}^*\mathbb{R}$ shares a similar approach to the one above. We will start by defining an equivalence relation $\sim$ on the set of all real sequences S and define ${}^*\mathbb{R}$ to be the set of equivalence classes $S/\sim$. Subsequently we lift the addition and multiplication operations, as well as the total order componentwise from the reals to the set of nonstandard reals.

When constructing the reals from the rationals, we saw that the goal was to create a Dedekind-complete extension of the rationals. This forms the main motivation behind the definition of its equivalence relation. When constructing the nonstandard reals however, Dedekind-completion is no longer the goal. In fact, we don't want the set of nonstandard reals to be Dedekind-complete at all. After all, if it were, it would be a Dedekind-complete ordered field and therefore isomorphic to the real numbers. It wouldn't be a 'true' extension of $\mathbb{R}$ anymore; something we would like ${}^*\mathbb{R}$ to be. Losing Dedekind-completeness is somewhat unfortunate, but there does exist a lesser version in the form of Theorem 1.2.2, which we'll get to later.

What *do* we want our set of nonstandard reals to be if not Dedekind-complete? It turns out we can express this through its elements, the equivalence classes.

Firstly we want ${}^*\mathbb{R}$ to be rich enough to allow for new elements such as the infinitesimals. For this reason we don't want the equivalence classes to be too large. For example, we'd like sequences that converge at

[1] See for example [2, p. 21].

different rates to be part of different equivalence classes; something which, when constructing the reals from the rationals, was not the case. Under the equivalence relation $\equiv$ for example, we see that the sequences $\{\frac{1}{n}\}$ and $\{\frac{1}{\sqrt{n}}\}$ are part of the same equivalence class; while in $^*\mathbb{R}$ we'd like them to be in different ones.

On the other hand, we don't want the equivalence classes to be too small such that $^*\mathbb{R}$ loses its interesting algebraic properties. The most extreme case can be seen in the following example.

Example 1.1.1. Let $\{a_n\} \sim \{b_n\}$ whenever $a_n = b_n$ for all n. In this case our equivalence classes are as small as possible. If we define multiplication in the most obvious way, namely compontentwise, we end up with the situation that

$$[\{1, 0, 1, 0, \ldots\}] \cdot [\{0, 1, 0, 1, \ldots\}] = [0].$$

Since neither of the elements on the left are equal to the equivalence class [0], we've ended up in the situation where there are zero divisors. Critically, $^*\mathbb{R}$ can no longer be a field.

We want the size of the equivalence classes to be just large enough to avoid the problem of zero divisors. To this end we will use a modified version of the equivalence relation of the previous example. Instead of calling two sequence equivalent when the components are equal everywhere, we will call them equivalent if the components are equal *almost* everywhere.

To formalize this concept, we need the following theorem:

Theorem 1.1.2. *There exists a function $\mu : \mathbb{N} \to \{0, 1\}$, such that for all $N, M \subseteq \mathbb{N}$*

(i) *μ is finitely additive: if $N \cap M = \emptyset$, then $\mu(N \uplus M) = \mu(N) + \mu(M)$,*

(ii) *$\mu(\mathbb{N}) = 1$ and $\mu(\emptyset) = 0$,*

(iii) *if $N \subseteq \mathbb{N}$ is finite, then $\mu(N) = 0$.*

Proof. We will proof the existence of such a function in Theorem 2.1.1 of Section 2.1. The theorem is, among other things, and exercise in Zorn's Lemma. □

As will become more clear in a moment, will refer to such a function μ as a *finitely additive measure*. Similarly for all subsets $N \subseteq \mathbb{N}$ we will call $\mu(N)$ the *measure* of N.

Like other proofs of existence that rely on Zorn's Lemma, this proof does not give an explicit expression for that which is proved to exist. Furthermore, it does not guarantee uniqueness of the measure. As we will only use the above properties however, which all such finitely additive measure share, this will not be a problem. From here on out we therefore fix μ to always refer to the same finitely additive measure that has the above properties.

Corollary 1.1.3. *Let N and M be subsets of $\mathbb{N}$.*

(i) *μ is strongly additive: $\mu(N \cup M) + \mu(N \cap M) = \mu(N) + \mu(M)$.*

(ii) *μ is subadditive: $\mu(N \cup M) \leq \mu(N) + \mu(M)$.*

(iii) *If $\mu(N) = 1$ and $N \subseteq M$, then $\mu(M) = 1$.*

(iv) *We either have $\mu(N) = 1$ or $\mu(N^c) = 1$, but not both.*

(v) *If $\mu(N) = 1$ and $\mu(M) = 1$, then $\mu(N \cap M) = 1$.*

Proof. (i) We can write $N \cup M$ as the disjoint union $N \uplus (M - N \cap M)$. Therefore by repeated application of finite additivity

$$\begin{aligned}\mu(N \cup M) + \mu(N \cap N) &= \mu(N \uplus (M - N \cap M)) + \mu(N \cap M),\\ &= \mu(N) + \mu(M - N \cap M) + \mu(N \cap M),\\ &= \mu(N) + \mu(M).\end{aligned}$$

(ii) First of all, note that $\mu(N \cap M) \geq 0$, therefore $\mu(N \cup M) \leq \mu(N \cup M) + \mu(N \cap M)$. If we now apply strong additivity on the right hand side we get $\mu(N \cup M) \leq \mu(N) + \mu(M)$.

(iii) If $N \subseteq M$, then we can write $M = N \uplus (M - N)$. Since μ is finitely additive, we have that

$$\mu(M) = \mu(N) + \mu(M - N).$$

But since $\mu(N) = 1$ by assumption, we see that $\mu(M - N)$ has to have value 0. Otherwise $\mu(M) = 2$ which is not it the range of μ. Therefore $\mu(M) = 1 + 0 = 1$.

(iv) For every subset of N of $\mathbb{N}$ we can write $\mathbb{N} = N \uplus N^c$. By finite additivity we have

$$\mu(N) + \mu(N^c) = \mu(N \uplus N^c) = \mu(\mathbb{N}) = 1.$$

Since the total has to equal 1, we see that we can rule out the possibilities of N and N^c both having measure 0, as well as both having measure 1. Therefore either $\mu(N) = 1$ or $\mu(N^c) = 0$, but not both.

(v) Consider the set $(N \cap M)^c$, we know this equals the set $N^c \cup M^c$. By subadditivity we have that

$$\mu(N^c \cup M^c) \leq \mu(N^c) + \mu(M^c).$$

Since $\mu(N) = 1$ and $\mu(M) = 1$, we know by the previous result (iv), that the sum on the right is zero, so the left has to be as well. Therefore by finite additivity

$$\mu(N \cap M) = \mu(N \cap M) + \mu(N^c \cup M^c) = \mu((N \cap M) \uplus (N \cap M)^c) = \mu(\mathbb{N}) = 1.$$

□

In general when dealing with a measure space $(X, \mathcal{A}, \nu)$, the phrase "φ ν-almost everywhere" means that the set of all $x \in X$ such that the statement φ is ***not*** true, is contained in a ν-null set. This is a set $N \in \mathcal{A}$ of which the measure $\nu(N)$ is zero. In other words: "φ ν-almost everywhere" means that there is an $N \in \mathcal{A}$ such that

$$\{x \in X \mid \varphi \text{ is not true}\} \subseteq N \text{ and } \nu(N) = 0.$$

Applying this terminology in our case, we see that $X = \mathbb{N}$, $\mathcal{A} = \mathcal{P}(\mathbb{N})$ and $\nu = \mu$. Since here the σ-algebra is just the power set of $\mathbb{N}$, we do not need to bother with the technicality that "φ μ-almost everywhere" means that the set $\{n \in \mathbb{N} \mid \varphi \text{ is true}\}$ *is contained* in a μ-null set; since in this case every subset of $\mathbb{N}$ is measurable so $\{n \in \mathbb{N} \mid \varphi \text{ is true}\}$ *is* a μ-null set. Additionally, since we only deal with one measure in the rest of this chapter, we drop the prefix "$\mu-$" from now on.

Using Corollary 1.1.3(iv), we see that if φ almost everywhere then the set $\{n \in N \mid \phi \text{ is true}\}$ has to have measure one. We see that in our case, showing that φ almost everywhere, is equivalent to showing that the subset of the natural numbers on which φ is true has measure one. So

$$\mu\{n \in \mathbb{N} \mid \varphi \text{ is not true}\} = 0, \text{ precisely when } \mu\{n \in \mathbb{N} \mid \varphi \text{ is true}\} = 1.$$

A lot of the interesting behaviour of μ is explained by combining its binary nature with its finite additivity. This also leads to results that can seem arbitrary, for example:

Example 1.1.4. Let E be the set of all natural numbers that are even, and O the set of all natural numbers that are odd. From Corollary 1.1.3(iv) follows that either $\mu(E) = 1$ or $\mu(E^c) = \mu(O) = 1$, but not both. Since we don't know the exact expression for μ, we don't know which one is which. So even though the sets E and O are isomorphic, and in that way contain the same amount of elements, the measure μ assigns them a different value.

With this measure μ, we are now ready to define the set of nonstandard reals ${}^*\mathbb{R}$:

Definition 1.1.5. Let S be the set of all real sequences. We define $\sim$ to be the relation on S, such that

$$\{a_n\} \sim \{b_n\}, \text{ precisely when } a_n = b_n \text{ almost everywhere.}$$

The resulting set of equivalence classes $S/\sim$ will be denoted by ${}^*\mathbb{R}$, and its elements are often called *hyperreals*, but in this text *nonstandard reals*. If $[a_n]$ denotes the equivalence class of $\{a_n\}$, *addition* and *multiplication* can be defined as follows

$$[a_n] + [b_n] = [a_n + b_n] \text{ and } [a_n] \cdot [b_n] = [a_n \cdot b_n].$$

Finally we *order* ${}^*\mathbb{R}$ such that

$$[a_n] < [b_n] \text{ precisely when } a_n < b_n \text{ almost everywhere.}$$

Before showing that ${}^*\mathbb{R}$ is an ordered field, we need to show the following book-keeping lemmas:

Lemma 1.1.6. *The relation $\sim$ is an equivalence relation.*

Proof. We need to check that $\sim$ is reflexive, symmetric and transitive.

- (*Reflexivity*): For all sequences of real numbers $\{a_n\}$ it follows that $a_n = a_n$ everywhere.
- (*Transitivity*): Suppose that $\{a_n\} \sim \{b_n\}$ and $\{b_n\} \sim \{c_n\}$. By definition of $\sim$, there are subsets N and M of $\mathbb{N}$ such that $\mu(N) = \mu(M) = 1$ and $a_n = b_n$ for all $n \in N$ while $b_n = c_n$ for all $n \in M$. Since $a_n = c_n$ when $n \in N \cap M$, we have that
$$N \cap M \subseteq \{n \in \mathbb{N} \mid a_n = b_n\}.$$
Because of Corollary 1.1.3(v), it follows that $\mu(N \cap M) = 1$. By Corollary 1.1.3(iii) and the above inclusion that means that $\{n \in \mathbb{N} \mid a_n = b_n\}$ also has measure one. Therefore $a_n = b_n$ almost everywhere.
- (*Symmetry*): When $a_n = b_n$ almost everywhere, then also $b_n = a_n$ almost everywhere.

□

Lemma 1.1.7. *The definitions for addition, multiplication and order are well defined.*

Proof. We need to show that these definitions are independent of the representatives of $[a_n]$ and $[b_n]$. Since the proofs are very similar, we we only show the most difficult of the three: the one for the order.

Suppose both $\{a_n\}$ and $\{A_n\}$ are representatives of $[a_n]$, and similarly $\{b_n\}$ and $\{B_n\}$ are representatives of $[b_n]$. Then there are sets N and M such that $\mu(N) = \mu(M) = 1 =$ and $a_n = A_n$ for all $n \in N$, and $b_n = B_n$ for all $n \in M$.

Suppose that $a_n < b_n$ almost everywhere as well. There exists a set U such that $\mu(U) = 1$, and $a_n < b_n$ when $n \in U$. If we take n from the triple intersection $N \cap M \cap U$, we see that
$$A_n = a_n < b_n = B_n.$$
By repeated application of 1.1.3(v), we have that $\mu(N \cap M \cap U) = 1$. So $A_n < B_n$ almost everywhere. □

So far we've seen a common element in the proof of both lemmas: the use of 1.1.3(v). In fact, when proving the algebraic properties of $^*\mathbb{R}$ the application of this result is so universal, that most of these proves look like reskinned versions of each other. For that reason I will leave the proof of the following theorem and corollary to you.

Theorem 1.1.8. *The set of nonstandard reals $^*\mathbb{R}$ is a ordered field.*

Corollary 1.1.9. *The map that sends all real numbers r, to the equivalence class $[r]$ is an order preserving ring homomorphism between the reals and the nonstandard reals.*

Importantly, we are now equipped with all the operations and expected behaviour that ordered fields have. Notably, we are able to consider inverse elements of nonstandard reals, and absolute value is defined as you would expect.

Instead of giving the proofs of the above results, we will look back on the problem of zero divisors and see that it is solved by the properties of μ.

Example 1.1.10. Let $\{a_n\} = \{1, 0, 1, 0, \ldots\}$ and $\{b_n\} = \{0, 1, 0, 1, \ldots\}$. From Example 1.1.4 we know that either $a_n = 0$ and $b_n = 1$ almost everywhere, or the opposite. In both cases it follows that
$$[\{1, 0, 1, 0, \ldots\}] \cdot [\{1, 0, 1, 0, \ldots\}] = [0] \cdot [1] = 0.$$

We now have the ordered field $^*\mathbb{R}$ that extends $\mathbb{R}$. The next step is to identify what sort elements $^*\mathbb{R}$ contains. We will do so by using the already familiar elements: the embedded reals.

Definition 1.1.11.
- An element $x \in {}^*\mathbb{R}$ is called *infinitesimal* when $|x| < r$ for all positive real numbers $r \in \mathbb{R}_{>0}$.
- An element $x \in {}^*\mathbb{R}$ is called *finite* when there is a positive real number $r \in \mathbb{R}_{>0}$, such that $|x| < r$. An element that is not finite is called *infinite*.

The following are some examples of these nonstandard elements.

Example 1.1.12. The element $[\frac{1}{n}]$ is infinitesimal. To show this, we need to prove $[\frac{1}{n}] < r$ for all positive real numbers $r \in \mathbb{R}_{>0}$. This comes down to comparing the sequences
$$\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\} \text{ and } \{r, r, r, r, \ldots\} \text{ for some } r \in \mathbb{R}_{>0}.$$

Since the sequence $\{\frac{1}{n}\}$ converges to zero, we know that for any real number r it holds that $\frac{1}{n} \geq r$ for only finite amount of n. Therefore $\{n \mid \frac{1}{n} \geq r\}$ has measure zero, and conversely $\frac{1}{n} < r$ almost everywhere.

Similarly, it can be shown that $[\frac{1}{\sqrt{n}}]$ is infinitesimal and that $[n]$ is infinite.

The following results can easily be verified.

Corollary 1.1.13. (i) *The sum of two infinitesimals is infinitesimal.*

(ii) *The product of a finite element with an infinitesimal, is again an infinitesimal.*

(iii) *There is but one infinitesimal that is real: the element zero.*

Proof. We will prove the second result only; the first follows in similar manner, while the third can be checked easily. Let x be infinitesimal and y finite. Since y is finite, there is some $R \in \mathbb{R}_{>0}$ such that $|y| < R$. And since x is infinitesimal, for any positive real number r follows

$$-r < x < r, \text{ so } -ry < xy < ry \text{ and } -rR < xy < rR.$$

Since rR can be made into an arbitrary positive real number, xy is again infinitesimal. □

One can wonder if these definitions cover all possible nonstandard elements. It turns out: not exactly. In case of the finite elements we instead see the following:

Proposition 1.1.14. *Any finite nonstandard element x can be written uniquely as a sum $x = r + \epsilon$, where r is a real number, and ϵ is infinitesimal.*

Proof. For proving the existence of r and ϵ, we will exploit the Dedekind-completeness of the real numbers. Let $r = \sup\{a \in \mathbb{R} \mid a < x\}$. To avoid confusion: note that while we take the supremum of a subset of the *real* numbers, we have defined this set using the *nonstandard* relation for order. Since x is finite, there exists a real upper bound, so a least upper bound r is well defined.

We want to show that $x - r$ is infinitesimal. Suppose it is not. Then there is a positive real number b such that $|x - r| > b > 0$. We consider the following cases

- When $x - r > 0$, it follows that $x > b + r > r$. Since $b + r$ is real, this contradicts the fact that r is an upper bound of the set $\{a \in \mathbb{R} \mid a < x\}$.
- When $x - r < 0$, we have that $r - x > b > 0$, so $r > x + b > x$. Since $x + b$ is real, this contradicts the fact that r is the *least* upper bound of the set $\{a \in \mathbb{R} \mid a < x\}$.

If we let $\epsilon := x - r$, we get the existence of a sum such that $x = r + \epsilon$.

Uniqueness of the sum is easier to proof: if $x = r_1 + \epsilon_1 = r_2 + \epsilon_2$, then $r_1 - r_2 = \epsilon_1 - \epsilon_2$. Since the element of the left is real, and the one on the right infinitesimal, by Corollary 1.1.13(iii) it follows that both sides must be zero. □

We will end this section with introducing one last set of new terminology, as well as a final result.

Definition 1.1.15. We say that two nonstandard elements x and y are *infinitely close*, when $x - y$ is infinitesimal and we write $x \approx y$.

From Proposition 1.1.14 follows that for each finite nonstandard element x, there is a unique real number r, such that $x \approx r$. We call this r the *standard part* of x and denote it by $°x$ or $\mathrm{st}(x)$. Conversely, for each $r \in \mathbb{R}$, the set of all nonstandard real elements x such that $\mathrm{st}(x) = r$ is called the *monad* of r.

In example Example 1.1.12 we already saw that the nonstandard elements associated with real sequences converging to zero, are infinitesimal. The following proposition is a generalization of that result.

Proposition 1.1.16. *If a real sequence $\{a_n\}$ has limit a, then the nonstandard element $[a_n]$ is infinitely close to a. So $[a_n] \approx a$.*

Proof. If the real sequence $\{a_n\}$ has limit a, then the sequence $\{a_n - a\}$ has limit 0. Therefore, applying the same technique we used in Example 1.1.12, we get that $[a_n - a] = [a_n] - [a]$ is infinitesimal. So $[a_n] \approx a$. □

We now have our mathematical structure we set out to construct: an ordered field that extends the real numbers, which includes the infinitesimal elements. In the next section we will discuss some of the interesting properties this new structure brings us.

1.2 Internal Sets and Functions

Now that we know what elements the set of nonstandard reals contains, we can start looking at its properties. As is usual, when presented with a new mathematical structure, we identify which objects we consider nice to work with. Think of open sets in topology, or measurable sets in measure theory. In our case we will call these objects ‘internal’. We have both internal sets and internal functions.

Definition 1.2.1. Let $\{A_n\}$ be a sequence of subsets of $\mathbb{R}$. This sequence defines a subset $[A_n]$ of $^*\mathbb{R}$ as follows:

$$[A_n] := \{[x_n] \in {}^*R \mid x_n \in A_n \text{ almost everywhere}\}.$$

When A is a subset of $^*\mathbb{R}$ and there is such an associated sequence of real subsets, then we call A an *internal set*. Similarly, if $\{f_n\}$ is a sequence of functions $f_n : \mathbb{R} \to \mathbb{R}$, then this defines the function $[f_n] : {}^*\mathbb{R} \to {}^*\mathbb{R}$ such that for all nonstandard reals $[x_n]$

$$[f_n]([x_n]) := [f_n(x_n)].$$

If a function $f : {}^*\mathbb{R} \to {}^*\mathbb{R}$ has such an associated sequence of real functions, then f is called an *internal function*. Objects that are not internal we call *external*.

What makes these objects nice to work with, is that they, similar to the nonstandard reals themselves, have an associated sequence of objects (either sets or functions) in $\mathbb{R}$. This makes it easy to show that certain properties hold for internal objects. We can simply lift them from their associated representatives in $\mathbb{R}$. For example:

Theorem 1.2.2 (Nonstandard Dedekind-Completeness)**.** *Let A be a nonempty internal subset of $^*\mathbb{R}$. If A is bounded from above, then it has a least upper bound.*

Proof. Since A is internal it has an associated sequence of real subsets $\{A_n\}$ and we can write $A = [A_n]$. Additionally, there exists some $[R_n] \in {}^*\mathbb{R}$ that is an upper bound of $[A_n]$. Combining the two, we see that for any element $[x_n] \in A$ we have both $x_n \in A_n$ and $x_n \leq R_n$ almost everywhere. Therefore A_n is bounded from above almost anywhere. Using the Dedekind-completeness of the real numbers the least upper bound $\sup A_n$ then also exists almost anywhere. We can then define the sequence $\{a_n\}$ such that $a_n = \sup A_n$ when the latter is defined, and $a_n = 0$ when its not. By this definition it must follow that $[x_n] \leq [a_n]$, and that it is the least upper bound of A. □

Remark 1.2.3. In the proof of the above result we have used made a lot of implicit use of Corollary 1.1.3(v). For example: by definition of internal sets $[x_n] \in [A_n]$ means that there is a set $N \subseteq \mathbb{N}$ with measure 1 such that $x_n \in A_n$ when $n \in N$. Additionally, by definition of the nonstandard order $[x_n] \leq [R_n]$ means that there is a set $M \subseteq \mathbb{N}$ with measure 1 such that $x_n \leq R_n$ when $n \in M$. Therefore both these properties hold when $n \in N \cap M$. By Corollary 1.1.3(v) the intersection $N \cap M$ also has measure 1 and we say that $x_n \in A_n$ and $x_n \leq R_n$ almost everywhere. As this idea is always the same, we won’t explicitly mention the use of Corollary 1.1.3(v) anymore.

A special type of internal sets and functions are the *standard* ones. These are the internal sets where their associated sequences are constant. That is to say:

Definition 1.2.4. If a subset $^*\mathbb{R}$ is of the form $[A]$ for some real subset $A \subseteq \mathbb{R}$, then we call $[A]$ *standard* and denote it by $^*A$. Similarly, if $^*f : {}^*\mathbb{R} \to {}^*\mathbb{R}$ is of the form $[f]$ for some $f : \mathbb{R} \to \mathbb{R}$ then it is also called *standard*.

Somewhat confusingly however we will, depending on context, also refer to $^*A$ and $^*f$ as the *nonstandard versions* of A and f.

Remark 1.2.5. These apparent contradictory definitions, are a question of perspective. When we want to emphasize that we had a real subset $A \subseteq \mathbb{R}$ and are now considering its nonstandard ‘equivalent’ $^*A$, we call the latter the nonstandard version of A. If however we start with a nonstandard set $S \subseteq {}^*\mathbb{R}$ and see that it can be written as $S = {}^*A$ for some real subset A, then we say that it is a standard set.

While we intuitively identify A with its nonstandard version $^*A$, in general the latter is much more rich.

Example 1.2.6. Consider the nonstandard natural numbers $^*\mathbb{N}$. By definition it contains all $[N_n]$ such that $N_n \in \mathbb{N}$. So it contains all embedded natural numbers in the form $[k] = [\{k, k, k, \ldots\}]$ for some $k \in \mathbb{N}$. But also the element $[n] = [\{0, 1, 2, \ldots\}]$ which is infinite.

This new terminology also allows us to look at 'nonstandard' versions of existing properties. Take the Archimedean criterion for example. In its standard form it would certainly fail to hold in the nonstandard reals. Take any infinite element of $^*\mathbb{R}$ and by its definition there can not be an embedded natural number that is larger. However, if we somewhat expand the property to allow for nonstandard natural numbers, it does hold:

Theorem 1.2.7 (Nonstandard Archimedean Property)**.** *For all $R \in {}^*\mathbb{R}$, there exist a nonstandard natural number $N \in {}^*\mathbb{N}$ such that $R < N$.*

Proof. We can write $R = [r_n]$ where $\{r_n\}$ is some real valued sequence. Since the real numbers are Archimedean, we have the existence of natural numbers N_n such that $r_n < N_n$ for all $n \in \mathbb{N}$. Therefore $[r_n] < [N_n]$ (almost) everywhere. □

The Nonstandard Construction

2.1 Ultrafilters

Looking back at our construction of the nonstandard reals, we see that it ultimately hinges on the existence and behaviour of the finitely additive measure we labeled μ. Back in the first chapter, we simply posited this measure but here we will show its existence.

The central result of this section will be the following theorem.

Theorem 2.1.1. *Let I be some set and U a collection of subsets of I. The function $\mu : \mathcal{P}(I) \to \{0, 1\}$ such that for all $A, B \subseteq I$*

$$\mu(A) = \begin{cases} 1 \text{ if } A \in U, \\ 0 \text{ if } A \notin U, \end{cases}$$

has the following properties if and only if U is a nonprincipal ultrafilter:

(M1) *$\mu(I) = 1$ and $\mu(\emptyset) = 0$;*

(M2) *If $\mu(A) = 1$ and $\mu(B) = 1$, then $\mu(A \cap B) = 1$;*

(M3) *If $\mu(A) = 1$ and $A \subseteq B$, then $\mu(B) = 1$;*

(M4) *Either $\mu(A) = 1$ or $\mu(I - A) = 1$;*

(M5) *If $A \cap B = \emptyset$ then $\mu(A \cup B) = \mu(A) + \mu(B)$;*

(M6) *If A is finite, then $\mu(A) = 0$.*

While this is quite the list of properties, they should come as no surprise as we have seen and used similar behaviour of μ back in the first chapter. What is new however, is the object U known as a 'nonprincipal ultrafilter'. As the above theorem indicates, correctly defining U give us our desired function μ. Central to this section, is therefore the theory of filters and ultrafilters.

Definition 2.1.2 (Filter)**.** Let I be a set. A collection U of subsets of I is called a *filter* on I if

(F1) $I \in U$ and $\emptyset \notin U$,

(F2) U is closed under finite intersection, so $A, B \in U$ implies $A \cap B \in U$,

(F3) U is upwardly closed, so $A \in U$ and $A \subseteq B \subseteq I$ imply $B \in U$.

When U is a filter on I, the latter is sometimes called an *index*.

Example 2.1.3. An immediate example of a filter on a nonempty set I, is the collection $U = \{I\}$. This is called the *trivial* filter on I.

Unsurprisingly trivial filters are of little interest. For the purpose of arriving at some nontrivial filters there is a useful property known as the *finite intersection property*. A collection S is said to have this trait if for all $A, B \in S$ the intersection $A \cap B$ is nonempty. The usefulness of this characteristic is demonstrated by the following lemma.

Lemma 2.1.4. *Let I be a set and S a collection of subsets of I.*

(i) *If S has the finite intersection property, then*

$$G_S := \{P \subseteq I \mid \text{ there are } T_0, \dots T_n \in S \text{ for some } n \in \mathbb{N} \text{ such that } T_0 \cap \dots \cap T_n \subseteq P\}$$

is the least filter on I containing S with regards to the inclusion relation.

(ii) *The set S is contained in a filter U on I if and only if S has the finite intersection property.*

Proof. (i) That G_S is a filter that contains S can be easily checked. To show that it is also the least such filter, suppose that there is another filter U on I such that $S \subseteq U \subset G_S$. Then there is $A \in (G_S - U)$ and $T_0, \dots, T_n \in S$ such that $T_0 \cap \dots \cap T_n \subseteq A$ for some $n \in \mathbb{N}$. However since U is closed under finite intersection and U contains S, we have that $T_0 \cap \dots \cap T_n \in U$. And since U is upwardly closed and $T_0 \cap \dots \cap T_n \subseteq A$ we also have $A \in U$. This contradicts our assumption that $A \in (G_S - U)$. Therefore $(G_S - U)$ is empty, and G_S is the minimal filter on I containing S.

(ii) By the previous result we already know that if S has the finite intersection property, it is contained in the filter G_S. Conversely: suppose that S is contained in a filter U. Then U has the finite intersection property since it is closed under finite intersection and does not contain $\emptyset$. Therefore any subset of U, in this case S, has to have the finite intersection property as well.

□

The previous result lets us extend any collection S that has the finite intersection property into a filter G_S. An important subcase is when S contains but one set. We can write $S = \{A\}$ for some $A \subseteq I$ and see that this type of collection always has the finite intersection property as long as $A \neq \emptyset$. The associated filter $G_{\{A\}}$ then takes the form

$$G_{\{A\}} = \{P \subseteq I \mid \text{ there is } T \in \{A\} \text{ such that } T \subseteq P\}.$$

So the least filter containing $\{A\}$ is the collection of all subsets of I containing A. As this is an important subcase we will be somewhat sloppy and write G_A instead of $G_{\{A\}}$ whenever A is an nonempty subset of I. Additionally we will introduce some new terminology:

Definition 2.1.5. Let I be a set and U a filter on I. If there is a subset A of I such that $U = G_A$, then we will call U a *principal* filter on I. We will also say that U is *generated* by A.

One type of principal filters on I are all filters on I containing a singleton; that is $\{x\}$ for some $x \in I$. We will see that these are automatically principal.

Lemma 2.1.6. *Let I be a set and U a filter on I. If U contains some singleton $\{x\}$ with $x \in I$, then U is generated by $\{x\}$. We can write $U = G_{\{x\}}$.*

Proof. If U contains the singleton $\{x\}$, then all elements of U have to contain $\{x\}$. If not, then we would have an empty intersection somewhere. We see that $U \subseteq G_{\{x\}}$. On the other hand, since $\{x\} \in U$ and U is upwardly closed, we see that $G_{\{x\}} \subseteq U$. Therefore $U = G_{\{x\}}$ and U is principal. □

There is an inherent link between the principal filters and the sets that generate them. This allows us to construct new filters as the following lemma shows.

Lemma 2.1.7. *Let I be a set and U a principal filter on I which generated by $A \subseteq I$. If V is a filter on I containing U, then the collection*

$$V \cap A := \{P \cap A \mid P \in V\}$$

is a filter on A contained in V.

Proof. Before we check that $V \cap A$ has the properties of a filter, note that since U is generated by A, we have that $A \in U \subseteq V$. And since V is closed under finite intersection, it follows that $V \cap A \subseteq V$.

(F1) Since $I \in V$, we see that $I \cap A = A \in U \cap A$. Furthermore, since $V \cap A$ is a collection of sets contained in the filter V, by Lemma 2.1.4(ii) it has the finite intersection property. So $\emptyset \notin V \cap A$.

(F2) Let $Q_1, Q_2 \in V \cap A \subseteq V$. Since V is closed under finite intersection we have $Q_1 \cap Q_2 \in V$. By definition both Q_1 and Q_2 are contained in A. So we see that $Q_1 \cap Q_2 \cap A = Q_1 \cap Q_2 \in V \cap A$.

(F3) Let $Q \in V \cap A \subseteq V$ and $C \subseteq A$ such that Q is contained in C. Since V is upwardly closed we have $C \in V$. Therefore $C \cap A = C \in V \cap A$.

□

The importance of principal filters will become clear later. For now, let us look back at Theorem 2.1.1 and the function μ we are trying try create. Comparing properties (M1-3) to the properties of a filter, (F1-3), we see that they are taken care of if we let U be a filter on I. The next property, (M4), is a much stronger statement however, for which we will need an equally strong concept. This will come in the form of collections known as 'ultrafilters'.

Definition 2.1.8 (Ultrafilter)**.** Let I be a set and P the family of all filters on I. Combined with the inclusion relation, this family P is a partially ordered set, or poset, $(P, \subseteq)$. We will call maximal elements of this poset *ultrafilters*. In other words, if U is an ultrafilter on I then for any other filter V on I such that $U \subseteq V$, we have $U = V$.

Similar to when we defined the concept of a filters, it will be useful to know when a collection S can be extended into an ultrafilter. The answer turns out to be the same. However, this time we are not able to give the explicit form the ultrafilter takes as it is an exercise in Zorn's Lemma. The following theorem demonstrates.

Theorem 2.1.9. *Let I be a set and U a filter on I. Then there is an ultrafilter V that contains U.*

Proof. We will use Zorn's Lemma to show the existence of V. Let Q be the family of all filters on I containing U. This set is a poset when combined with the inclusion relation. It is also nonempty since U is itself an element of Q. We now consider a chain C of Q. In other words, C is a totally ordered subset of Q. We can write $C := \{A_n \mid n \in N\}$ for some totally ordered set N and we have $A_n \subseteq A_m$ whenever $n \leq m$.

We have to show that C has an upper bound A. That is to say, A has to be a filter that contains all elements of the chain. If C is the empty chain, then any element of Q is an upper bound of C. This element exist since Q is nonempty. If C is not empty, then it is not hard to show that $A := \bigcup_{n \in N} A_n$ is a filter. By definition of A, we also have that $A_n \subseteq A$ for any $n \in N$. Therefore A is an upper bound of C.

We now have shown that an arbitrary chain C has an upper bound. Hence any chain of $(Q, \subseteq)$ has an upper bound. Invoking Zorn's Lemma, we get that Q contains a maximal element V. Since Q consisted of all filters containing U, this maximal element V is an ultrafilter containing U. □

Corollary 2.1.10. *Let I be a set. If S is a collection of subsets of I that has the finite intersection property, then there is an ultrafilter U on I that contains S.*

Proof. This is simply a result of combining the above theorem with Lemma 2.1.4. □

The reason we are interested in ultrafilters is that through their maximality, ultrafilters gain a lot of powerful properties.

Lemma 2.1.11. *Let I be a set and U an ultrafilter on I. For all $A, B \subseteq I$*

(i) *either $A \in U$ or $(I - A) \in U$, but not both,*

(ii) *if $A \cup B \in U$ then $A \in U$ or $B \in U$.*

Proof. (i) We will exclude both the possibility of U containing A and $(I - A)$, and the possibility of U containing neither.

Firstly, if U contains A and $(I - A)$, then since U is closed under finite intersections we have that $A \cap (I - A) = \emptyset \in U$, contradicting U being a filter. Secondly, if U contains neither A nor $(I - A)$, then we can separate two cases:

- For all $S \in U$ we have $A \cap S \neq \emptyset$. Then $U \cup \{A\}$ has the finite intersection property and by Lemma 2.1.4 it can be extended into a filter V strictly larger than U. This contradicts the maximality of U;
- For some $S \in U$ we have $A \cap S = \emptyset$. It follows that $S \subseteq (I - A)$. Since U is upwardly closed we have $(I - A) \in U$ which contradicts our assumption.

We see that U has to contain either A or $(I - A)$, but can't contain both.

(ii) Suppose that $A \cup B \in U$ and that U contains neither A nor B. By the property we just showed, we know that $(I - A) \in U$ and $(I - B) \in U$. Since U is closed under finite intersection, the triple intersection $(A \cup B) \cap (I - A) \cap (I - B)$ is contained in U. However, this intersection is empty contradicting U being a filter. Therefore if U contains $A \cup B$, it also contains A or B.

□

The result of Lemma 2.1.11(i) can actually be used as an alternative definition for ultrafilters as the following corollary shows.

Corollary 2.1.12. *Let I be a set and U a filter on I. Then U is an ultrafilter if and only if $A \in U$ or $(I - A) \in U$ for all $A \subseteq I$.*

Proof. Lemma 2.1.11(i) has already shown the left to right implication. The reverse implication is clear: if U is a filter and it contains either A or $(I - A)$ for all $A \subseteq I$, it can't actually contain more sets while remaining a filter. If we try to add any new element $B \subseteq U$, we see that the 'complement' $(I - B)$ is already contained. Adding this element would result into having to add their intersection, which is empty. So U is maximal with regards to the inclusion relation and therefore an ultrafilter. □

We see that if μ is defined using an ultrafilter U on I, then according to Lemma 2.1.11(i) it has property (M4). In fact, the same lemma shows the finite additivity of μ, property (M5), as well:

Corollary 2.1.13. *Let I be a set and U an ultrafilter on I. The function $\mu : \mathcal{P}(I) \to \{0, 1\}$ such that* *for all* $A \subseteq I$

$$\mu(A) = \begin{cases} 1 \text{ if } A \in U, \\ 0 \text{ if } A \notin U, \end{cases}$$

is finitely additive.

Proof. We need to show that for all disjoint $A, B \subseteq I$ the following equation holds:

$$\mu(A \uplus B) = \mu(A) + \mu(B).$$

Since U is a filter, it can't contain both A and B. Otherwise their empty intersection would also be included in U. Therefore there are only two cases to consider:

- Either A or B is an element of U, but not both. In which case we take $A \in U$ and $B \notin U$. Then, since $A \subseteq A \uplus B$ and because U is upwardly closed, we get $A \uplus B \in U$. So we see that $\mu(A \uplus B) = 1$, $\mu(A) = 1$ and $\mu(B) = 0$. Substituting into the above equation we get $1 = 1$;
- Neither A nor B is an element of U. By Lemma 2.1.11(ii) we see that $A \uplus B$ is also not included in U. Therefore $\mu(A \uplus B) = 0$, $\mu(A) = 0$ and $\mu(B) = 0$. The above equation states $0 = 0$.

We see that in both cases μ is indeed finitely additive. □

We are left with the final property, (M6). Sticking to our ultrafilter approach, we somehow need to make sure that the ultrafilter U does not contain any finite sets. It is here we see the return of principal filters. Recall that we referred to a filter U on I as principal, if there was a subset $A \subseteq I$ such that $U = G_A$. The following theorem will seem unrelated to our goals. However, as its corollary will show, it demonstrates exactly what type of extra condition we have to place on our ultrafilter.

Theorem 2.1.14. *Let I be a set. If U is a principal filter on I generated by $A \subseteq I$, then there is a bijection between the ultrafilters on I containing U and the ultrafilters on A.*

Before we start with its proof, let us show the following lemma.

Lemma 2.1.15. *Let I be a set and U a principal filter on I generated by A.*

(i) *If V is an ultrafilter on I containing U, then $V \cap A$ is an ultrafilter on A.*

(ii) *If W is an ultrafilter on A, then $G_{W \cup U}$ is an ultrafilter on I containing U.*

Proof. (i) By Lemma 2.1.7 we already know that $V \cap A$ is a filter on A contained in V. What is left, is to show that it is also a maximal filter on A. We will do so by using the alternative characteristic for ultrafilters from Corollary 2.1.12.

Let $R \subseteq A$. Since V is an ultrafilter on I we have $R \in V$ or $(I - R) \in V$. Then by definition either $R \cap A = R \in V \cap A$ or $(I - R) \cap A \in V \cap A$. In case of the latter we see that $(I - R) \cap A$ can be rewritten as $(A - R)$. So either $R \in V \cap A$ or $(A - R) \in V \cap A$. Applying Corollary 2.1.12 we see that $V \cap A$ is an ultrafilter on A.

(ii) Since U is generated by A, all elements of the former contain the latter. Combined with the fact that all elements of W are contained in A, we see that, for all $P \in W$ and $Q \in U$ we can write $P \cap Q$ as $P \cap A$ instead. Since A is nonempty if it generates a filter, and W has the finite intersection property, we have that $P \cap A$ is nonempty. So the collection $W \cup U$ has the finite intersection property as well. Applying Lemma 2.1.4(i) we see that $G_{W \cup U}$ is a filter on I containing $W \cup U$. To show that it is also an ultrafilter, we will once again use the alternative characteristic of ultrafilters.

Let $R \subseteq I$. Then, since $R \cap A \subseteq A$ and W is an ultrafilter on A, we have that $R \cap A \in W$ or $A - (R \cap A) \in W$. This second case can be rewritten as $A - (R \cap A) = A - R$. So either $R \cap A \in W$ or $(A - R) \in W$. Since $G_{W \cup U}$ is upwardly closed we see that in the first case $R \cap A \subseteq R$, so $R \in G_{W \cup U}$. Similarly, in the second case we see $(A - R) \subseteq (I - R)$, so $(I - R) \in G_{W \cup U}$. By Corollary 2.1.12 we have that $G_{W \cup U}$ is an ultrafilter on I.

□

We return to the proof of Theorem 2.1.14. Using the previous lemma we can explicitly construct the desired bijection.

Proof. (Theorem 2.1.14) Let P be the family of all ultrafilters on I that contain U, and Q the family of all ultrafilters on A. Then the function $f : P \to Q$ such that $f(V) = V \cap A$ for all $V \in P$ is well defined by Lemma 2.1.15. Furthermore it is both injective and surjective as we will show.

- (*Injectivity*) Suppose that $V_1 \cap A = V_2 \cap A$ but $V_1 \neq V_2$. We take $V_2 - V_1$ to be nonempty and let $R \in V_2 - V_1$. By assumption there is an $S \in V_1$ such that $R \cap A = S \cap A$. But then we have $S \cap A \subseteq R$. And since V_1 is upwardly closed we have $R \in V_1$ which contradicts our choice of R.
- (*Surjectivity*) We will prove that $G_{W \cup U} \cap A = W$ for all $W \in Q$ by showing that inclusion holds both ways.

  Firstly, if we take $R \in G_{W \cup U} \cap A$ then by definition R can be written as $S \cap A$ where there are $T_0, \ldots, T_n \in W \cup U$ such that $T_0 \cap \ldots \cap T_n \subseteq S$. We will consider the finite intersection $D := T_0 \cap \ldots \cap T_n \cap A$. Note that if any $T_i \in U$ then they contain A, so they actually do not contribute to this intersection. And if all $T_i \in U$ then $D = A \in W$. Without loss of generality we can therefore assume that all $T_0, \ldots, T_n \in W$. In which case D is just a finite intersection of elements of W, and therefore also included in W. We see that $D \subseteq S \cap A = R$, and since $D \in W$ and W is upwardly closed, we have $R \in W$.

  Secondly, if we take $R \in W$, we automatically see that $W \in G_{W \cup U}$ since the latter is a filter containing $W \cup U$.

We see that $f : P \to Q$ is a bijection with $f^{-1}(W) = G_{W \cup U}$ for all $W \in Q$. □

While this result is seemingly unrelated to our goals, it turns out to be precisely what we need as the following corollary demonstrates.

Corollary 2.1.16. *Let I be a set and U an ultrafilter on I.*

(i) *U is principal if and only if it is generated by a singleton.*

(ii) *If U contains any finite sets, then it contains a singleton.*

Proof. (i) If U is generated by the singleton $\{x\}$, it by definition principal. We are left with the reverse implication; suppose that U is a principal ultrafilter on I generated by some subset A of I. This means that the set of all ultrafilters containing U contains but one item: U itself. By Theorem 2.1.14 there is a bijection from this set to all ultrafilters on A. Therefore there is also but one ultrafilter on A.

If A is not a singleton, then there are at least two different $x, y \in A$. By Corollary 2.1.10 there are also two ultrafilters W_1 and W_2 on A containing either $\{x\}$ or $\{y\}$. These ultrafilters can not be equal since then they would have to contain both $\{x\}$ and $\{y\}$ and therefore their empty intersection as well. This contradicts there only being one ultrafilter on A. Therefore A can't contain more than one element.

On the other hand, since A generates the filter U, it can not be empty either. So A contains exactly one element, and can be written as $\{x\}$ for some $x \in I$.

(ii) Suppose that the ultrafilter U contains a finite set. For this finite set we can write $A_n := \{a_0, \ldots, a_n\}$ for some $n \in \mathbb{N}$. We will describe a process with which we will eventually arrive at a singleton contained in U:

- If $n = 0$ then $A_0 = \{a_0\} \in U$ and we are done.
- If $n > 0$ then we can consider $A_{n-1} = \{a_0, \ldots, a_{n-1}\}$. Since U is an ultrafilter we see that $A_{n-1} \in U$ or $(I - A_{n-1}) \in U$. In the second case we have $A_n \cap (I - A_{n-1}) = \{a_n\} \in U$ and we are also done. In the first case we can repeat the process for A_{n-1}.

Since A_n is finite, it is clear that this process will end. Therefore U contains $\{a_i\}$ for some $i \in \{0, \ldots, n\}$.

□

Combining these results with Lemma 2.1.6, we see that if U is a principal ultrafilter, then U contains a singleton, which is a finite set. Conversely, if U is an ultrafilter that contains a finite set, it contains a singleton and is therefore principal. Hence the ultrafilter U contains finite elements if and only if it is principal. The following theorem should therefore come as no surprise.

Theorem 2.1.17. *Let I be a set and U an ultrafilter on I. Then U contains no finite sets if and only if it is nonprincipal.*

Proof. The result is a direct application of Corollary 2.1.16 combined with Lemma 2.1.6. □

An obvious question would be if nonprincipal ultrafilter actually exist. On a finite index I they clearly do not. So a necessary requirement is for I to be infinite. If U is to be a nonprincipal ultrafilter on an infinite I, then if A is a finite subset of I, it is not contained in U by the previous theorem. Therefore it does contain $(I - A)$, which is a cofinite set. Hence the nonprincipal ultrafilter U contains the collection of all cofinite sets on I. This is useful, since this collection is in fact a filter itself:

Example 2.1.18. Let I be an infinite set. The collection of all cofinite sets F, is a filter on I: the so called *Fréchet Filter*.

That F contains the infinite I and does not contain the finite $\emptyset$ is immediate. It is also obvious that F is upwardly closed. To see that it is also closed under finite intersections, note that for all $A, B \in F$ we have that $I - (A \cap B) \subseteq (I - A) \cup (I - B)$. Since the latter union is finite, the former is as well. Therefore $A \cap B \in F$.

We can now show that nonprincipal ultrafilters do exist on infinite indices: we can simply extend Fréchet Filters into ultrafilters and show these have to be nonprincipal. The following theorem makes this explicit.

Theorem 2.1.19. *Let I be an infinite set. Then there exists a nonprincipal ultrafilter on I.*

Proof. By Example 2.1.18, the Fréchet Filter on I is indeed a filter. Therefore there is an ultrafilter extending it by Theorem 2.1.9. This extension can't contain any finite sets, as their complement, the cofinite sets, are already included. The ultrafilter is therefore principal. □

Looking back at the theorem we set out to prove, we now see that we have all the tools we need:

Proof. (Theorem 2.1.1) Recall that the theorem was set up both ways: if $\mu : \mathcal{P}(I) \to \{0, 1\}$ is defined as

$$\mu(A) = \begin{cases} 1 \text{ if } A \in U, \\ 0 \text{ if } A \notin U, \end{cases}$$

then it has properties (M1) to (M6) if and only if U is a nonprincipal ultrafilter on I.

We see that properties (M1) to (M3) are satisfied precisely when U is a filter on I, while property (M4) is the alternative characterication of ultrafilters from Lemma 2.1.11(i). So the first four properties are satisfied when U is an ultrafilter on I. The finite additivity in the form of property (M5) is a direct consequence of U being an ultrafilter as we learned through Lemma 2.1.11(ii). Finally, the last property, (M6), implies that U does not contain finite sets. Combined with the other properties and Theorem 2.1.17, this is satisfied if and only if U is also nonprincipal. □

We see that back in the first chapter we hid a lot of the theory of filters and ultrafilters behind the measure μ. This was done as to not detract from that its main purpose: getting the reader up to speed with concepts from nonstandard analysis through its most familiar application, the nonstandard reals. In keeping with this mindset, the main takeaway of this sections should be the existence and behaviour of the measure μ. Similar to some theorems of other fields such as the Carathéodory Extention Theorem, the importance is that it *can* be done and not so much *how* it is done. While ultrafilters are certainly a powerful tool, applied in many fields of mathematics, they are, in a way, only tangentially related to nonstandard analysis. More on this later.

As a final remark of this section, you might wonder why we settled for a finitely additive measure and not for a proper measure that is σ-additive. Well, the answer is that in most cases it's the best we can do. See the following example.

Example 2.1.20. Let U be a nonprincipal ultrafilter on $\mathbb{N}$. We know it exists since $\mathbb{N}$ is infinite. We also know that it contains no singletons $\{n\}$ for some $n \in \mathbb{N}$. Therefore $\mu(\{n\}) = 0$. On the other hand we have that $\mathbb{N} \in U$, so $\mu(\mathbb{N}) = 1$.

When we consider the pairwise disjoint sequence of sets $(\{n\})_{n \geq 0}$, we see that their disjoint union $\biguplus_{n \geq 0}\{n\}$ is $\mathbb{N}$. However

$$\mu(\mathbb{N}) = 1 \neq 0 = \sum_{n \geq 0} \mu(\{n\})$$

Therefore μ on the index $\mathbb{N}$ is not σ-additive.

2.2 Ultraproducts

Equipped with the finitely additive measure μ of Theorem 2.1.1, we can use it to characterise the nonstandard version of any set S with infinite index I. As we will see, the nonstandard version ${}^*S$ will be a construct more generally known as an *ultrapower*.

For the purposes of nonstandard analysis, we are primarily interested in the case that $I = \mathbb{N}$. However, the theory of ultrapowers, and its more general theory of *ultraproducts*, does not get significantly more difficult when considering arbitrary infinite I. Therefore we will not restrict ourselves to the natural numbers. Instead we will treat the more general theory, and then apply it to our setting at the end.

As a starting point for the construction of the nonstandard reals we considered the set of all real sequences. Since we now work with a more general index, we also need a more general version of this concept. A suitable candidate is the following.

Definition 2.2.1. Let I be a set and have X_i be a set for each $i \in I$. Then the *product* $\prod_{i\in I} X_i$ is the set of all functions $f : I \to \bigcup_{i\in I} X_i$ such that $f(i) \in X_i$ for each $i \in I$.

Remark 2.2.2. The reason for calling this type of set a product becomes clear when we let $I = \{0, \dots, n\}$ for some $n \in \mathbb{N}$ and have $X_i = \mathbb{R}$ for all $i \in I$. A function of the type $f : \{0, \dots, n\} \to \mathbb{R}$ can be identified with the $(n + 1)$-tuple $(f(0), \dots, f(n))$ in the Cartesian product $\mathbb{R}^{n+1}$. Conversely, every $(n + 1)$-tuple of the Cartesian product defines such a function in $\prod_{i\in I} X_i$.

In the case that $I = \mathbb{N}$ and $X_i = S$, we see that the product $\prod_{i\in\mathbb{N}} S$ is in fact the set of all S-valued sequences. Therefore the concept of a product properly extends the concept of S-valued sequences. It will therefore serve as our new starting point for our nonstandard construction.

The next step is defining an equivalence relation on this product.

Theorem 2.2.3. *Let I be a set and have X_i be a set for all $i \in I$. Let $p, q \in \prod_{i\in I} X_i$. If U is a filter on I then the following relation is an equivalence relation:*

$$p \sim_U q \text{ if and only if } \{i \in I \mid p(i) = q(i)\} \in U.$$

Proof. The relation is clearly symmetric. We are left with showing that it is also reflexive and transitive. The first follows from the fact that since U is a filter, we have $I \in U$. The latter follows from the other two properties of a filter, as is shown below.

Let $p \sim_U q$ and $q \sim_U r$. Then by definition there are $A, B \in U$ such that

$$A := \{i \in I \mid p(i) = q(i)\} \text{ and } B := \{i \in I \mid q(i) = r(i)\}.$$

Since U is a filter, it is closed under finite intersection, as well as upwardly closed. By the first property we have that $A \cap B \in U$. So for all $i \in A \cap B$ it follows that $p(i) = r(i)$. Therefore

$$A \cap B \subseteq \{i \in I \mid p(i) = r(i)\}.$$

Being upwardly closed, we see that U has to contain $\{i \in I \mid p(i) = r(i)\}$. Therefore $p \sim_U r$ and the relation is transitive. □

Our interest now shifts to the quotient space $\prod_{i\in I} X_i/\sim_U$. An important case is when U is also an ultrafilter on I.

Definition 2.2.4. Let I be a set and have X_i be a set for all $i \in I$. If U is an ultrafilter on I we call the quotient space $\prod_{i\in I} X_i/\sim_U$ an *ultraproduct*. We will denote it by $\prod_U X_i$. When all X_i are equal we can write $\prod_U S$ for some set S. In this case we call it an *ultrapower* instead.

The concept of an ultrapower is sufficient for our purposes and we will us it to introduce some nonstandard-analysis-specific notation.

Definition 2.2.5. Let S be a set. Let U be a nonprincipal ultrafilter on $\mathbb{N}$. Then we will call the ultraproduct $\prod_U S$ a *nonstandard version* of S and denote it by ${}^*S_U$.

By the previous section there is an alternative way to characterise nonstandard versions of sets. Namely, for every nonprincipal ultrafilter U there is an associated finitely additive measure μ_U in the form of Theorem 2.1.1. This measure was defined such that $A \in U$ if and only if $\mu_U(A) = 1$. Therefore, looking back at the equivalence relation in Theorem 2.2.3, we see that for all $p, q \in \prod_{n\in\mathbb{N}} S$

$$p \sim_U q \text{ if and only if } \mu_U(\{i \in I \mid p(i) = q(i)\}) = 1.$$

At this point we can reintroduce our previous terminology and rephrase this statement as

$$p \sim_U q \text{ if and only if } p(i) = q(i) \text{ almost everywhere.}$$

So ${}^*S_U$ is the set of equivalence classes under the above relation.

An apparent problem now however, is that there are (possibly) multiple nonstandard versions of any set S. This is especially troubling seeing as in the first chapter, we referred to ${}^*\mathbb{R}$ as *the* nonstandard reals. Recall however, that at the start of that chapter we picked and then subsequently fixed μ as some measure having the properties in Theorem 1.1.2. We now know through Theorem 2.1.1 that this is equivalent to picking some nonprincipal ultrafilter U. Therefore, the set we constructed and called the nonstandard reals, is actually a nonstandard version of the real numbers denoted by ${}^*\mathbb{R}_U$.

While this does mean that everything in the first chapter is well defined, it still leaves the question if any of its results are dependent on the choice of U. The answer is 'no'. All the results were derived from the properties of $\mathbb{R}$ and the properties of Theorem 1.1.2 which itself was a reformulation of Theorem 2.1.1. The latter are shared between all nonprincipal ultrafilters, and therefore any result derived from these two sources hold in all nonstandard versions of $\mathbb{R}$.

Similarly, as long as we restrict our derivations on any nonstandard version of S to the properties of S as well as the properties of all nonprinciple ultrafilters on $\mathbb{N}$ outlined in Theorem 1.1.2, the subsequent results will hold in all nonstandard versions of S.

Because of this reason, whenever we talk about *the* nonstandard version of a set S, we refer to any of its nonstandard versions: simply pick one. To formalise this concept we will introduce the **-transform* (to be read as *nonstandard transform*) which maps any set S to one of its nonstandard versions. We will denote the resulting nonstandard version by ${}^*S$.

We can show that this *-transform exists since by Theorem 2.1.19 there exists *a* nonprincipal ultrafilter on $\mathbb{N}$. Therefore any set S has *a* nonstandard version. We then invoke the Axiom of Choice to pick which is *the* nonstandard version. As we already used an Axiom of Choice equivalent in the form of Zorn's Lemma to show that ultrafilters actually exist, this is not a some new concession on our part.

With this new terminology, we can 'overwrite' the previous definition of the nonstandard reals ${}^*\mathbb{R}_U$ by the *-transform of $\mathbb{R}$. We see that for this new ${}^*\mathbb{R}$, all the results of the first chapter still hold.

In a more general setting, we can now transform any set S into its nonstandard version ${}^*S$. The elements of ${}^*S$ are the equivalence classes of the relation $\sim_U$ for some nonprincipal ultrafilter U on $\mathbb{N}$. This relation has an equivalent characterisation through the use of the finitely additive measure μ_U from Theorem 2.1.1 and the use of the 'almost everywhere' terminology. It is this characterisation we will use from now on.

The Nonstandard Elements

Since we now have the ability to transform any set into its nonstandard version, we once again would like to be able to identify which nonstandard objects are nice to work with. In the first chapter we saw two examples of such objects: the internal sets and the internal functions.

Recall that the internal set $[A_n]$ referred to the set that contained all nonstandard reals $[x_n]$ such that $x_n \in A_n$ almost everywhere. Similarly, the internal function $[f_n]$ referred to the function such that $[f_n]([x_n]) = [f_n(x_n)]$.

When we reflect on this 'internal' label, we see that its defining characteristic is that these internal objects have an associated sequence of either sets or functions in the real numbers. This is exactly what made these objects pleasant to work with, as we were able to transfer properties that hold for either sets or functions in the real numbers, to their associated internal object. For example, the nonstandard version of Dedekind-completion Theorem 1.2.2 holds for all internal sets through an argument that can be summarised as "it holds for all their associated sets in the real numbers." In this sense these internal objects mirror the nonstandard reals themselves. The latter have by definition an associated sequence of reals through one of their representatives of their equivalence class.

Note that we can extend this notion of internal much further than just sets and functions, as well as not just limit it to the nonstandard reals. We would like to be able to refer to any nonstandard object on any nonstandard set ${}^*S$ as 'internal' as long as it has an associated sequence of the 'same' objects on the standard set S. Internal functionals for instance; these would be functionals $[F_n]$ of ${}^*S$ defined on internal functions $[f_n]$ of ${}^*S$ such that $[F_n]([f_n]) = [F_n(f_n)]$. In a similar vein, we would also like to talk about 'internal sets of internal functionals' and in general be able to use the label 'internal' for any type of object on ${}^*S$.

In this section we will cover all of these internal objects at once by introducing the concept of *internal elements*. To do this, we will first introduce the concept of *superstructures*. The idea behind this construction is that a superstructure of S, in a way, contains all information of S. This includes its sets, functions and functionals etc. By considering this superstructure, we are then able to discuss all of these objects at once, and therefore able to more easily define a concept of internal which covers all of them.

3.1 Superstructures

Definition 3.1.1 (Superstructure). Let S be a set. For all $n \in \mathbb{N}$ will define the *levels* $V_n(S)$ of S inductively by

$$V_0(S) = S \text{ and } V_{n+1}(S) = V_n(S) \cup \mathcal{P}(V_n(S)).$$

We will call the union $\bigcup_{n\in\mathbb{N}} V_n(S)$ of all levels of S the *superstructure* over S, and denote it by $V(S)$. The *V-transform*, to be read as *superstructure transform*, is the function that maps each set S to its superstructure $V(S)$.

Corollary 3.1.2. *Let S be a set, $V(S)$ its superstructure and $n, m \in \mathbb{N}$.*

(i) *The n-th level is contained in all equal or larger levels: so if $n \leq m$, then $V_n(S) \subseteq V_m(S)$.*

(ii) *The n-th level is an element of all strictly larger levels: so if $n < m$, then $V_n(S) \in V_m(S)$.*

Proof. By definition of $V_n(S)$ we have that $V_n(S) \subseteq V_{n+1}(S)$ for all $n \in \mathbb{N}$. So for all $n \leq m$ it follows that

$$V_n \subseteq V_{n+1} \subseteq \ldots V_{m-1} \subseteq V_m.$$

Since inclusion is transitive, we see that $V_n \subseteq V_m$ for all $n \leq m$, and this shows (i).

For the second property, notice that by definition it also follows that $V_n(S) \in V_{n+1}(S)$. Therefore, whenever $n < m$ we have by the first part (i) that $V_{n+1}(S) \subseteq V_m(S)$. So $V_n(S) \in V_{n+1}(S) \subseteq V_m(S)$, and (ii) follows. □

The central idea behind a superstructure over a set S, is that we order all set-theoretical constructs on S through its levels $V_n(S)$. Correspondingly, the *rank* of an element x in $V(S)$ is the least $n \in \mathbb{N}$ such that $x \in V_n(S)$. We will write $\text{rank}_S(x) = n$. Whenever the superstructure $V(S)$ is clear from context we will drop the subscript 'S' and write $\text{rank}(x) = n$ instead.

Lemma 3.1.3. *Let S be a set and $V(S)$ its superstructure. If $x \in V(S)$ has rank $n \in \mathbb{N}$, then*

(i) $x \in V_m(S)$ *for all* $m \geq n$,

(ii) $x \subseteq V_{n-1}(S)$ *if* $n > 0$,

(iii) $x \subseteq V(S)$ *if* $n > 0$.

Proof. (i) Since $\operatorname{rank}(x) = n$, we have that $x \in V_n(S)$. By Corollary 3.1.2(i) we know that the n-th level $V_n(S)$ is contained in all equal or larger levels. Therefore $x \in V_m(S)$ whenever $m \geq n$.

(ii) Suppose that $\operatorname{rank}(x) = n > 0$. By the minimality of the rank we have that x is not an element of all smaller levels than n. Notably x is not an element of $V_{n-1}(S)$. Since the n-th level is defined as $V_n(S) = V_{n-1}(S) \cup \mathcal{P}(V_{n-1}(S))$, we see that x has to be strictly contained in $\mathcal{P}(V_{n-1}(S))$. It is therefore a subset of the $(n-1)$-th level $V_{n-1}(S)$.

(iii) By the previous result, we have that x is a subset of the $(n-1)$-th level $V_{n-1}(S)$. Seeing as the superstructure was the union over all levels, x is also a subset of $V(S)$. So $x \subseteq V(S)$.

□

The third result can be rephrased as:

An element x of a superstructure $V(S)$, is either an element of S, or a subset of the superstructure.

By virtue of their definition, we can show that the levels of a given a set S have certain properties using induction. For example, we will call a set S *transitive* if and only if $S \subseteq \mathcal{P}(S)$.[1] We can show that if S is transitive, so are all its levels.

Lemma 3.1.4. *Let S be a transitive set. Then for all $n \in \mathbb{N}$*

(i) $V_n(S)$ *is transitive,*

(ii) $V_{n+1}(S) = \mathcal{P}(V_n(S))$.

Proof. (i) We will use induction on the levels as promised. If $n = 0$ we see that $V_0(S) = S$ is transitive by assumption. Therefore, suppose that $n > 0$ and that $V_n(S)$ is transitive. We need to show that $V_{n+1}(S) \subseteq \mathcal{P}(V_{n+1}(S))$.

Take an element $x \in V_{n+1}(S)$. Since $V_n(S)$ is transitive, we have that $V_n(S) \subseteq \mathcal{P}(V_n(S))$. By definition of the $(n+1)$-th level we also have $V_{n+1}(S) = V_n(S) \cup \mathcal{P}(V_n(S))$. This can therefore be written as $V_{n+1}(S) = \mathcal{P}(V_n(S))$. We see that an element of $V_{n+1}(S)$ is a subset of $V_n(S)$. So x has to be a subset of $V_n(S)$. By Corollary 3.1.2(i) it follows that the latter is contained in $V_{n+1}(S)$, so

$$x \subseteq V_n(S) \subseteq V_{n+1}(S).$$

Therefore an arbitrary element x of $V_{n+1}(S)$, is a subset of that same level. This is precisely what we needed to show, and we have $V_{n+1}(S) \subseteq \mathcal{P}(V_{n+1}(S))$. Therefore $V_{n+1}(S)$ is transitive.

(ii) We already implicitly showed this result proving the previous. For completeness sake however: by the first property all levels are transitive, therefore $V_n(S) \subseteq \mathcal{P}(V_n(S))$ for all $n \in \mathbb{N}$. By the definition of levels $V_{n+1}(S) = V_n(S) \cup \mathcal{P}(V_n(S))$. So this can be written as $V_{n+1}(S) = \mathcal{P}(V_n(S))$

□

A set that is trivially transitive is the empty one. While itself is empty, its superstructure is not.

Example 3.1.5. (Empty Set) Let $S = \emptyset$. Using Lemma 3.1.4, we know the levels of S are transitive and $V_{n+1}(\emptyset) = \mathcal{P}(V_n(\emptyset))$. Those familiar with the cumulative hierarchy will recognise that the levels take the form of the sets associated with all finite ordinals (the natural numbers) and that the superstructure $V(S) = \bigcup_{n \in \mathbb{N}} V_n$ is the set associated with the limit ordinal $\mathbb{N}$.

Regardless of knowledge of the cumulative hierarchy, we can show that there exist elements in $V(\emptyset)$ of any rank $n > 0$. We will do this through the use of the common identification of the natural numbers with certain set-theoretic constructs on the empty set. Inductively we associate the natural numbers with these constructs by:

$$\langle 0 \rangle := \emptyset \text{ and } \langle n+1 \rangle := \langle n \rangle \cup \{\langle n \rangle\}.$$

[1] This inclusion is of course a strict one, as famously proven by Cantors diagonal argument. In the following arguments however, for simplicity's sake, we will use the subset-or-equal sign.

This results in

$\langle 0\rangle = \emptyset$;

$\langle 1\rangle = \{\emptyset\} = \{\langle 0\rangle\}$;

$\langle 2\rangle = \{\emptyset, \{\emptyset\}\} = \{\langle 0\rangle, \langle 1\rangle\}$;

$\langle 3\rangle = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} = \{\langle 0\rangle, \langle 1\rangle, \langle 2\rangle\}$;

$\vdots$

$\langle n\rangle = \{\langle 0\rangle, \ldots, \langle n-1\rangle\}$ etc.

Using induction on levels we can show that $\text{rank}\langle n\rangle = n+1$. For $n = 0$ we have that $\langle 0\rangle = \emptyset \subseteq V_0(\emptyset)$. So $\langle 0\rangle$ is a subset of the 0-th level, and is therefore itself an element of $V_1(\emptyset) = \mathcal{P}(\emptyset)$.

For $n > 0$, suppose that we have $\text{rank}\langle n\rangle = n+1$. By the definition of rank we have $\langle n\rangle \in V_{n+1}(\emptyset) = \mathcal{P}(V_n(\emptyset))$ so $\langle n\rangle \subseteq V_n(\emptyset)$. Because of the latter we also have $\{\langle n\rangle\} \subseteq \mathcal{P}(V_n(\emptyset))$. Now recall the definition of $\langle n+1\rangle$, we see

$$\langle n+1\rangle = \langle n\rangle \cup \{\langle n\rangle\} \subseteq V_n(\emptyset) \cup \mathcal{P}(V_n(\emptyset)) = V_{n+1}(\emptyset).$$

A subset of the $(n+1)$-th level is an element of the $(n+2)$-th level, so $\langle n+1\rangle \in V_{n+2}(\emptyset)$ and $\text{rank}\langle n+1\rangle \leq (n+2)$.

We now need to rule out the possibility of this rank being smaller than $(n+2)$. Note that we can write $\langle n+1\rangle$ as the set $\{\langle 0\rangle, \ldots, \langle n\rangle\}$. So if the rank of $\langle n+1\rangle$ is strictly smaller than $(n+2)$ we would have $\{\langle 0\rangle, \ldots, \langle n\rangle\} \in V_{k+2}(\emptyset) = \mathcal{P}(V_{k+1}(\emptyset))$ for some $k < n$. But then the $\langle n\rangle$ would be an element of $V_{k+1}(\emptyset)$ and $\text{rank}\langle n\rangle \leq (k+1) < (n+1)$ contradicting our assumption. Therefore $\text{rank}\langle n+1\rangle = (n+2)$

The main takeaway from this example should be that even the simplest superstructure $V(\emptyset)$ has elements we can point to, that are of a specific rank larger than 0.

Our next step is to see how superstructures of different sets are related to each other. With induction on levels, the following result is an obvious one.

Lemma 3.1.6. *If S and T are sets such that $S \subseteq T$, then $V_n(S) \subseteq V_n(T)$ for all $n \in \mathbb{N}$. In particular $V(S) \subseteq V(T)$*

Proof. We will show that $V_n(S) \subseteq V_n(T)$ for all $n \in \mathbb{N}$. When $n = 0$, then $V_0(S) = S \subseteq T = V_0(T)$ by assumption. Therefore, let $n > 0$ and assume that $V_n(S) \subseteq V_n(T)$ holds. Because of this assumption we also have $\mathcal{P}(V_n(S)) \subseteq \mathcal{P}(V_n(T))$. Combining these two inclusions we get

$$V_n(S) \cup \mathcal{P}(V_n(S)) \subseteq V_n(T) \cup \mathcal{P}(V_n(T)).$$

Since these are exactly the definitions of de $(n+1)$-th layer of S and T respectively, we have that $V_{n+1}(S) \subseteq V_{n+1}(T)$.

By induction we include that $V_n(S) \subseteq V_n(T)$ for all $n \in \mathbb{N}$. As a superstructure is the union over its levels, the same inclusion holds for $V(S)$ and $V(T)$, so $V(S) \subseteq V(T)$. □

Intuitively, we would like the opposite of the statement in Lemma 3.1.6 to be true as well, i.e. if $V(S) \subseteq V(T)$ then $S \subseteq T$. As of now this is not the case.

Example 3.1.7. Let T be a set and $S \subseteq \mathcal{P}(T)$. By induction we can prove that $V_n(S) \subseteq V_{n+1}(T)$. For $n = 0$ we have

$$V_0(S) \subseteq \mathcal{P}(T) \subseteq T \cup \mathcal{P}(T) = V_1(T) \tag{3.1}$$

Let $n > 0$ and assume that $V_n(S) \subseteq V_{n+1}(T)$. Then also $\mathcal{P}(V_n(S)) \subseteq \mathcal{P}(V_{n+1}(T))$. Therefore

$$V_n(S) \cup \mathcal{P}(V_n(S)) \subseteq V_{n+1}(T) \cup \mathcal{P}(V_{n+1}(T)).$$

Identifying the levels on both sides we see that $V_{n+1}(S) \subseteq V_{n+2}(T)$. By induction we can conclude that every n-th level of S is contained in the $(n+1)$-th level of T. We see that $V(S) \subseteq V(T)$. However, S is not a subset of T.

What causes the reverse implication of Lemma 3.1.6 to fail in this example is that the elements of set S are themselves subsets of T. We can remedy this fact by requiring that the elements of whatever set we take the superstructures over, are not sets themselves. At least, we won't treat them as such.

Instead, whenever we are working with the concept of superstructures over a set S, we will treat the elements of S as *urelements*. These objects are not sets, and therefore contain no elements. They are not the empty set either, as that still would imply that they are a set. Union and intersection of urelements is therefore not defined. Still allowed however, is creating sets with urelements as their elements. Because of this, the powerset of a set with urelements is properly defined and our definition of levels remains a valid one.

What we are in effect doing when using urelements is setting the baseline of our set-theoretic constructions on a set S as the set S and we ignore the underlying structure of its elements.

From here on out, whenever we deal with a superstructure $V(S)$ we will treat the elements of S as urelements. With this convention we can show that the reverse of Lemma 3.1.6 holds.

Theorem 3.1.8. *If S and T are sets, then $S \subseteq T$ if and only if $V(S) \subseteq V(T)$.*

Proof. The left to right implication was precisely the statement of Lemma 3.1.6. By treating the elements of S and T as urelements, we can show the converse.

Let $x \in S$, we want to show that $x \in T$. By assumption we have that $\text{rank}_S(x) = 0$ and x is an urelement. We see that if we allow $\text{rank}_T(x) > 0$, that this quickly leads to problems as a rank larger than 0 would imply x having elements. This contradicts our convention so $\text{rank}_T(x)$ has to be 0, which means that $x \in T$. We conclude $S \subseteq T$. □

Now that we know how superstructures relate, it is no surprise to see that $V(\emptyset) \subseteq V(S)$ for any set S. Similar to Example 3.1.5, we can therefore show that any superstructure $V(S)$ contain elements of rank $n+1$ for all $n \in \mathbb{N}$: precisely the elements $\langle n \rangle$ from before.

So far we have gotten an idea what superstructures look like, and now we will turn to their application. As previously stated: the central idea is that a superstructure over S orders the set-theoretic constructions over S. This is useful to us as we can think of relations and functions as exactly that. We will show how this is done.

Definition 3.1.9. Given two sets or urelements x and y, we will call the collection $\{\{x\}, \{x, y\}\}$ the *ordered pair* of x and y and we will denote it by (x, y). When S and T are sets as well, the *Cartesian product* $S \times T$ is the family

$$S \times T := \{(x, y) \mid x \in S \text{ and } y \in T\}.$$

A *relation* R from S to T is any subset of this Cartesian product. We call such a relation a *function* from S to T if $(x, y) \in R$ and $(x, z) \in R$ implies that $y = z$.

Because relations and functions are now defined as set-theoretical constructs, they show up in the superstructure. For example:

Example 3.1.10. Let S be a set and $A, B \subseteq S$. If R is a relation from A to B, then by definition we can write

$$R = \{(x, y) \mid x \in A \text{ and } y \in B\}.$$

We can 'unpack' this construction to see that $\text{rank}(R) = 3$. We start with the ordered pair (x, y), by definition this was the collection $\{\{x\}, \{x, y\}\}$ for all $x \in A$ and $y \in B$. Since $\{x\}$ and $\{x, y\}$ are themselves subsets of $A \subseteq S$ and $A \cup B \subseteq S$ respectively, we have that $\{x\} \in \mathcal{P}(S)$ and $\{x, y\} \in \mathcal{P}(S)$. By definition of levels we have $\text{rank}(\{x\}) = 1$ and $\text{rank}(\{x, y\}) = 1$. So the ordered pair (x, y) is a collection of elements of rank 1. Therefore it is itself an element of rank 2. Finally, since the relation R is the family of these ordered pairs, it is of rank 3.

We see that this example can be extended to a more general setting.

Theorem 3.1.11. *Let S be a set and $V(S)$ its superstructure. If A and B are sets with elements of rank n for some $n \in \mathbb{N}$, then a relation R of A to B is of rank $(n+3)$.*

Proof. The result is very similar to the previous example. We have

$$R = \{(x, y) \mid x \in A \text{ and } y \in B\}.$$

By assumption all $x \in A$ and $y \in B$ have rank n. Therefore $\{x\}$ and $\{x, y\}$ are of rank $(n+1)$. The ordered pair $\{\{x\}, \{x, y\}\}$ is then of rank $(n+2)$. Finally, a set of ordered pairs is then of rank $(n+3)$. □

The above result holds for relations, so it also holds for functions. Specifically we see that a functions with a domain contained in S, to an image contained in S are elements of the superstructure of rank 3. A function of these functions (a functional) is then an element of rank 6. And functions of functionals are then elements of rank 9 etc.

In superstructures we find our ability to treat all objects on S at once: they are all simply elements of the superstructure. We will use this fact on our nonstandard framework to identify the internal elements.

3.2 Internal Elements

We have done a lot of preparatory work to get to this section. So far we have defined two types of mathematical constructs: the nonstandard construction and the superstructure embodied by the *-transform and the V-transform respectively. In this section these two come together to allow us to assign the 'internal' label to any nonstandard object.

As a reminder of what we are looking for: let S be a set; we want to be able to denote objects of the nonstandard construction $^*S$ as 'internal' when they have an associated sequence of standard objects. This was to extend the notion of internal sets and internal functions. Similar to the sets and functions we want to be able to talk about internal functionals, internal sets of internal functionals etc.

The superstructure provides a suitable framework. We can consider the superstructure of the nonstandard version of S which with our notation is denoted by $V(^*S)$. As we have seen in the previous section, this superstructure $V(^*S)$ contains all set-theoretic information of $^*S$. This includes its functions and functionals. It is to the elements of $V(^*S)$ we want to assign the label 'internal'.

On the other hand, we can also reverse the order of the transforms: we can make S into its superstructure $V(S)$ first, and then consider its nonstandard version. This becomes $^*V(S)$ in our notation. Recall that at its core the nonstandard version of a set P is a specific ultraproduct; its elements being equivalence classes of the product $\prod_{n\in\mathbb{N}} P$ under the 'almost everywhere' relation. In the case of $^*V(S)$, its elements are equivalence classes with representatives in $\prod_{n\in\mathbb{N}} V(S)$. This product consists of all sequences of elements of $V(S)$. So the elements of $^*V(S)$ have an associated representative that is a sequence of standard elements in $V(S)$.

The elements of $^*V(S)$ therefore act as *proto-interal* elements. The reason for the 'proto' affix, is that $^*V(S)$ is *not* the set $V(^*S)$, the latter being the one we are ultimately interested in. In this section we will look at how we can utilize these proto-internal elements to identify the actual internal elements of $V(^*S)$. We will do so with a function that maps the so called *bounded* proto-internals to the internal elements of $V(^*S)$.

In Figure 3.1 there is a diagram of the steps we just discussed.

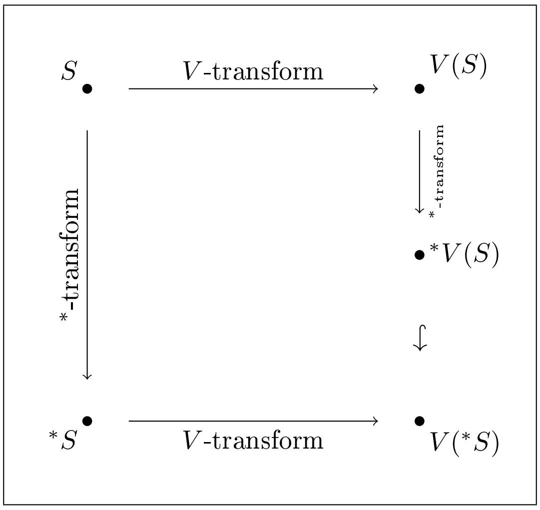

Figure 3.1: Given a set S, there are multiple ways to apply the *-transform and V-transform. We can consider both $V(^*S)$ and $^*V(S)$. In this section we will identify elements of the former using the elements of the latter, and call them 'internal'.

We will now formally begin with this identification process, starting with the elements of ${}^*V(S)$.

Definition 3.2.1. Let S be a set. We will call an element $x \in {}^*V(S)$ a *proto-internal element* of $V({}^*S)$. When $V({}^*S)$ is clear from context we sometimes call it a *proto-internal* instead.

While we are interested in proto-internals of $V({}^*S)$ because they have an associated sequence of elements in $V(S)$, not all of them are meaningful. Intuitively, if $[A_n]$ is a proto-internal and for all n we have $\text{rank}_{V(S)}(A_n) = k$ for some $k \in \mathbb{N}$, we want to associate it with an element of equal rank in $V({}^*S)$. For example, if a proto-internal has an associated sequence of functions, then we want to identify it with a function in $V({}^*S)$ as well. If however $[A_n]$ is a proto-internal and $\text{rank}(A_n) \neq \text{rank}(A_m)$ for all $n \neq m$, then there is not an obvious object in $V({}^*S)$ of the same rank to associate it with. We will therefore introduce the following condition on proto-internals.

Definition 3.2.2. Let S be a set. We will call a proto-internal $[A_n]$ of $V({}^*S)$ *bounded* if there is a $k \in \mathbb{N}$ such that $\text{rank}(A_n) \leq k$ for all $n \in \mathbb{N}$.

This bounded criterion is sufficient for our needs, as combined with the 'almost everywhere' terminology we get:

Lemma 3.2.3. *Let S be a set. If the proto-internal $[A_n]$ is bounded, then there is a unique natural number $k \in \mathbb{N}$ such that $\text{rank}(A_n) = k$ almost everywhere.*

Proof. If $[A_n]$ is bounded, then by definition there exists a $p \in \mathbb{N}$ such that

$$\mathbb{N} = \{n \mid \text{rank}(A_n) = 0\} \,\dot\cup\, \ldots \,\dot\cup\, \{n \mid \text{rank}(A_n) = p\}.$$

Since the entirety of the natural numbers has measure 1, and the union on the right is disjoint, there has to be a $0 \leq k \leq p$ such that

$$\mu\{n \mid \text{rank}(A_n) = k\} = 1.$$

Therefore $\text{rank}(A_n) = k$ almost everywhere. Uniqueness of this k follows quickly. For suppose there is $p \in \mathbb{N}$ such that $\text{rank}(A_n) = p$ almost everywhere as well. Then there are sets $A, B \subseteq \mathbb{N}$ such that $\{i \in A \mid \text{rank}(A_i) = k\}$ and $\{i \in B \mid \text{rank}(A_i) = p\}$ have measure 1. By Theorem 2.1.1 the intersection $A \cap B$ has measure 1 and is nonempty. Therefore when $i \in A \cap B$ then $k = \text{rank}(A_i) = p$. So we have $k = p$. □

Remark 3.2.4. We see a return of the 'almost everywhere' terminology. As this is the first time we use the convention outlined at the end of Section 2.2, we will be explicit in what this means:

We start with the set S and transform it into its superstructure $V(S)$. Then, we transform this into *the* nonstandard version of $V(S)$ which is ${}^*V(S)$. This nonstandard version has an associated nonprincipal ultrafilter U, and therefore there is a finitely additive measure μ_U such as in Theorem 2.1.1. It is with regards to this measure that 'almost everywhere' is defined. Since this ultrafilter U is fixed, so is μ_U. Therefore we will drop the U subscript and only write μ.

As a result of Lemma 3.2.3, we have that every bounded proto-internal of $V({}^*S)$ has an associated sequence of elements that are almost all of the same rank. In fact, we can show that it must have a representative whose elements all have the same rank.

Corollary 3.2.5. *Let S be a set and x be a bounded proto-internal on $V({}^*S)$. Then there exists a unique natural number $k \in \mathbb{N}$ and sequence $\{A_n\}$ of elements of $V(S)$ such that $\text{rank}(A_n) = k$ for all n and $x = [A_n]$.*

Proof. By Lemma 3.2.3 there exists a unique natural number k and a sequence $\{B_n\}$ such that $\text{rank}(B_n) = k$ almost everywhere. If $k = 0$ then this must mean that S is nonempty, so we can simply replace all B_n such that $\text{rank}(B_n) \neq k$ by any element of S.

If $k > 0$ then we know by the previous section that the objects we labeled $\langle k-1 \rangle$ from Example 3.1.5 are of rank k. In this case, simply replace all B_n such that $\text{rank}(B_n) \neq k$ by them instead.

In either case we can define a new set A_n that equals B_n when $\text{rank}(B_n) = k$ and equals its replacement when its not. Since by assumption $\text{rank}(B_n) = k$ almost everywhere, this replacement is only done on a set with measure 0. Therefore $x = [A_n]$. □

In this sense every bounded proto-internal $[A_n]$ of $V({}^*S)$ has a rank associated with it. We will therefore define the *rank of a bounded proto-internal* as the natural number k such that $\text{rank}(A_n) = k$ almost everywhere and we will write $\text{rank}[A_n] = k$. By Lemma 3.2.3 this number exists and is unique.

Importantly, we can now use induction on the rank of bounded proto-internals. It is this technique that we will use to define the internal elements of $V(^*S)$.

Definition 3.2.6. Let S be a set and let $\mathcal{B}$ denote the set of all bounded proto-internals. Let $[A_n] \in \mathcal{B}$. Using induction on the rank we will define function that sends each bounded proto-internal $[A_n]$ to an element $\overline{[A_n]}$ of $V(^*S)$.

If $\text{rank}[A_n] = 0$ then $A_n \in V_0(S) = S$ almost everywhere. By Corollary 3.2.5 we can take $A_n \in S$ *everywhere* without loss of generality. Since the elements $^*S$ are the equivalence classes on the S-valued sequences we can also consider $[A_n]$ as an element of $^*S$. It is therefore included in $V(^*S)$ as well. So we associate the proto-internal $[A_n]$ of rank 0, precisely with the element $[A_n]$ in $V(^*S)$. And we have $\overline{[A_n]} := [A_n]$ whenever $\text{rank}[A_n] = 0$.

If $\text{rank}[A_n] = n > 0$ then we define

$$\overline{[A_n]} := \{[B_n] \in \mathcal{B} \mid \text{rank}[B_n] < n \text{ and } B_n \in A_n \text{ almost everywhere}\}. \tag{3.2}$$

We will call an element of $V(^*S)$ that is of the form $\overline{[A_n]}$ for some bounded proto-internal $[A_n]$ an *internal element* of $V(^*S)$.

Appendix

A.1 Complete Ordered Fields

There are many different notions of a field being complete. The most common of which being a Cauchy-complete field, where all its Cauchy-sequences are also convergent in that field. In this section we will treat the notions of Cantor- and Dedekind-completeness, and how this relates to Cauchy-completeness. The final result of this section will be that every ordered field that is Dedekind-complete, is order-isomorphic to the field of the real numbers.

When dealing with multiple groups, rings and fields, it can become cumbersome to denote every operation and identity element with a different symbol. We have therefore elected not to do so in the following treatment. The symbols should be read contextually instead.

Definition A.1.1 (Ordered Field)**.** An *ordered field* is a field L with a total order $\leq$ that preserves the field operations. More precisely: for $a, b, c \in L$

(O_1) if $a < b$, then $a + c < b + c$,

(O_2) if $0 < a$ and $0 < b$, then $0 < a \cdot b$.

Corollary A.1.2. *Let L be an ordered field and $a, b, c \in L$. Then*

(i) if $a < b$ and $0 < c$, then $ac < bc$,

(ii) if L is non-trivial then $0 < 1$.

Proof. (i) If $a < b$, then by property (O_1), we can add $-a$ on both sides to get $0 < b - a$. If $0 < c$ as well, then we can use property (O_2) to get $0 < (b - a) \cdot c$. Using the distributive laws of L we get $0 < bc - ac$. Finally, we add ac on both sides using (O_1) to get $ac < bc$.

(ii) Let L be non-trivial. We will show $0 < 1$ by contradiction. Suppose that $1 \leq 0$. Since $0 = 1$ holds only in the trivial field, we have that $1 < 0$. So by property (O_1) follows that $1 + (-1) < 0 + (-1)$, so $0 < -1$. But then by property (O_1) we have that $0 < (-1) \cdot (-1) = 1$. So $1 < 0$ and $0 < 1$. Since $<$ is a (strict) total order this is a contradiction and the converse must hold.

□

Most of the expected behaviour of inequalities can be derived from the above results. For our purposes, the familiar results from the following lemma will be useful.

Lemma A.1.3. *Let L be a non-trivial ordered field and $a, b, c \in L$. Then*

(i) if $0 < c$ then $0 < \frac{1}{c}$,

(ii) if $0 < b, c$ and $\frac{a}{b} < c$ then $\frac{a}{c} < b$.

Proof. (i) Suppose that the converse is true, i.e. $\frac{1}{c} \leq 0$. Since c is positive, we can multiply both sides with c using Corollary A.1.2(i). We get $1 < 0$ which contradicts Corollary A.1.2(ii). Therefore $0 < \frac{1}{c}$.

(ii) By the first result we have that $0 < \frac{1}{c}$. So by property (O_2) we also get that $0 < \frac{b}{c}$. Finally using Corollary A.1.2(i), we can multiply both sides of the inequality $\frac{a}{b} < c$ with $\frac{c}{b}$, which gives the desired result.

□

For any field, there exists a unique function from the natural numbers to said field which maps 0 and 1 to the additive and multiplicative identity elements of the field. When the field is also non-trivial and ordered, this function is injective and embeds the natural numbers in the field. The following lemma makes this explicit.

Lemma A.1.4. *When L is a field, then there exists an unique function $i_L : \mathbb{N} \to L$, such that for all $n, m \in \mathbb{N}$*

(i) $i_L(1) = 1$,

(ii) $i_L(n + m) = i_L(n) + i_L(m)$.

If L is also non-trivial and ordered, then *(iii)* $i_L(n) < i_L(m)$ *when* $n < m$.

Proof. Both existence and uniqueness follow from combining the first two properties. Namely, if i_L is some function that has the first two properties, it would follow that $i_L(2) = 1 + 1$. So by induction we can show $i_L(n) = i_L(n-1) + 1$ for $n \in \mathbb{N}_{>1}$. Additionally, since $i_L(n+0) = i_L(n) + i_L(0)$, we see that $i_L(0) = 0$. So any function from the natural numbers to L that adheres to the first two properties, maps $n \in \mathbb{N}$ to the element that is n-times sum of $1 \in L$. Conversely, we get existence of such a function by defining $i_L(0) = 0$ and letting $i_L(n) = i_L(n) + 1$ for $n \in \mathbb{N}_{>0}$.

For the third property, we assume that L is a non-trivial ordered field. By Corollary A.1.2(ii) we have that $0 < 1$. By repeated application of property (O_1), we see that $i_L(n) < i_L(n) + 1 = i_L(n+1)$. Since the (strict) order is transitive, we have that $i_L(n) < i_L(m)$ for all $n, m \in \mathbb{N}$ such that $n < m$. □

Since for any given field L this function i_L is unique and, in the case of non-trivial ordered fields, it is an embedding of the natural numbers in L, we will from here on out say that the natural number n is an element of L. So for any natural number $n \in L$ we mean the element $i_L(n)$ in L. With this in mind, we can define the following property.

Definition A.1.5 (Archimedean Property)**.** An ordered field L is called *Archimedean* if for all elements of L, there exist a natural number whose embedded version is larger. In other words: for all $l \in L$ there exist $n \in \mathbb{N}$ such that $l < n$.

Remark A.1.6. An Archimedean ordered field is non-trivial since the Archimedean property requires at least a countably infinite amount of elements.

So far we've looked at ordered fields and some of their properties. We would now like to apply them. There is however, one last topic we need to look at first; namely what it means for a sequence to be convergent in an ordered field.

Strictly speaking: we will call a sequence convergent if it is so with respect to the order topology. In practice this means that we generalise the limit concept of metric spaces to include generalised metric spaces. We call a function a *generalised metric* if it has all the properties of a normal metric, except that it is no longer strictly real-valued. Instead it will be a L-valued metric when L is an ordered field. We do this since an ordered field is not necessarily a metric space. On the other hand: all ordered fields are generalised metric spaces through the use of the absolute value function. We formalize this idea with the following definitions.

Definition A.1.7 (Convergence)**.** Let L be an ordered field and $\{a_n\}$ a sequence in L. Let $\varepsilon \in L_{>0}$, then

(i) if there exist an $a \in L$ and $N \in \mathbb{N}$ such that for all $n \geq N$ it follows that $|a - a_n| < \varepsilon$, then we say that the sequence $\{a_n\}$ *converges* to a, and we write $\lim_{n\to\infty} a_n = a$,

(ii) if there exist an $N \in \mathbb{N}$ such that for all $n, m \geq N$ it follows that $|a_n - a_m| < \varepsilon$, then we call $\{a_n\}$ a *Cauchy-sequence*.

Remark A.1.8. When comparing these definitions to the limit concept of metric spaces, we see that ε is no longer a positive real number. Instead it is some positive element of L. Similarly, the role of the metric is replaced by a generalised one. In this case this is done through the absolute value function on L.

The following lemma is a familiar result in the ordered field of the real numbers. However, as the example following it shows, it is strictly a result of the Archimedean property.

Lemma A.1.9. *If L is an Archimedean ordered field, then*

$$\lim_{n\to\infty} \frac{1}{n} = 0.$$

Proof. Let $\varepsilon \in L_{>0}$. We need to find a $N \in \mathbb{N}$ such that $\frac{1}{n} < \varepsilon$ for all $n \geq N$. Since L is Archimedian, there exists a $N \in \mathbb{N}$ such that $\frac{1}{\varepsilon} < N$. So if $n \geq N$, then also $\frac{1}{\varepsilon} < n$. Rewriting this equation using Lemma A.1.3, we get that $\frac{1}{n} < \varepsilon$ for all $n \geq N$. Therefore the sequence $\{\frac{1}{n}\}$ converges to 0. □

Example A.1.10. The above result doesn't hold in case of the non-Archimedean field $^*\mathbb{R}$. If we let ε be some positive infinitesimal, then, since $\frac{1}{n}$ is real for all $n \in \mathbb{N}$, we have that $\varepsilon < \frac{1}{n}$ for all $n \in \mathbb{N}$, and the converse can never hold.

We are now ready to define three notions of completeness of an ordered field.

Definition A.1.11 (Completeness). Let L be an ordered field.

(i) (Dedekind-Complete) If all non-empty subsets of L that are bounded from above have a least upper bound, then we call it *Dedekind-Complete.*

(ii) (Cantor-Complete) If all nested sequences of non-empty closed intervals have a non-empty intersection, we call the field *Cantor-Complete.*

(iii) (Cauchy-Complete) If all Cauchy-sequences in L also converge in L, we call the field *Cauchy-complete.*

These notions of completeness are related as follows.

Theorem A.1.12 (Complete Ordered Fields). *The following statements are equivalent:*

(i) The ordered field L is Dedekind-complete.

(ii) The ordered field L is Cantor-complete and Archimedean.

(iii) The ordered field L is Cauchy-complete and Archimedean.

Before commencing with the proof, we will need one last result.

Lemma A.1.13. *Let L be a Cantor-complete ordered field. Suppose that $\{I_n\}$ is a nested sequence of non-empty closed intervals and the length of each interval $l(I_n)$ converges to 0. We write $I_n = [a_n, b_n]$ for each $n \in \mathbb{N}$. Then the sequences $\{a_n\}$ and $\{b_n\}$ converge to the same limit $x \in L$ and $\cap I_n = \{x\}$.*

Proof. We know that by Cantor-completess the intersection $\cap I_n$ is non-empty. Let $x, y \in \cap I_n$. Then for each $n \in \mathbb{N}$ we have that $x, y \in I_n$. Additionally for each $\varepsilon \in L_{>0}$, there is $N \in \mathbb{N}$ such that $l(I_n) = b_n - a_n < \varepsilon$ when $n \geq N$. So $|x - y| < b_n - a_n < \varepsilon$ for all $\varepsilon \in L_{>0}$ and sufficiently large n. Clearly $x = y$. So $\cap I_n$ has at most one element and since it's non-empty, it has exactly one element. Finally, since the interval I_n contains a_n and b_n as well as x for all $n \in \mathbb{N}$, and its length goes to zero for sufficiently large n, the sequences $\{a_n\}$ and $\{b_n\}$ both converge to x. □

We now return to the proof of the theorem:

Proof (Theorem A.1.12).

- $[(i) \implies (ii)]$: Let L be a Dedekind-complete ordered field. We need to show that every nested sequence of non-empty closed intervals has a non-empty intersection. Let $\{I_n\}$ be a such a sequence. We write $I_n = [a_n, b_n]$, for all $n \in \mathbb{N}$. Let $A := \{a_n \mid n \in \mathbb{N}\}$ and $B := \{b_n \mid n \in \mathbb{N}\}$, and accordingly $a := \sup(A)$ and $b := \sup(B)$. Since L is Dedekind-complete, a and b are well defined. Now since $\{I_n\}$ is a nested sequence, we have that $a_n \leq b_m$ for all $n, m \in \mathbb{N}$. So b_n is an upper bound of A and a_n is a lower bound of B for all $n \in \mathbb{N}$. It follows that $a_n \leq a \leq b \leq b_m$ for all $n, m \in \mathbb{N}$. Therefore the interval $[a, b]$ is contained in I_n for all $n \in \mathbb{N}$. So $[a, b] \subseteq \cap_{n\in\mathbb{N}} I_n$, and since the former is non-empty, the intersection is non-empty as well. This show the Cantor-completeness.

  We will show that Archimedean property follows from Dedekind-completeness as well. Assume the converse; then there exists a $l \in L$ such that $n \leq l$ for all $n \in \mathbb{N}$. So the set of natural numbers is bounded, and by Dedekind-completeness has a least upper bound a. It follows that $(a - 1)$ should also be an upper bound. If it were not, then $[(a - 1), a]$ would contain at least one natural number N; but then $a \leq (N + 1) \in \mathbb{N}$ and a would no longer be an upper bound of the natural numbers. However, $(a - 1)$ also being an upper bound contradicts the minimality of a. So if L is Dedekind-complete, it has to be Archimedean as well.

- $[(ii) \implies (iii)]$: Let L be an Archimedean Cantor-complete ordered field. We wish to show that every Cauchy-sequences in L converges in L. Let $\{a_n\}$ be a such a sequence. We will define a suitable sequence of intervals such that we can use Cantor-completeness. For all $p \in \mathbb{N}_{>0}$, let $N_p \in \mathbb{N}$ be the minimal natural number such that $|a_n - a_m| < (2p)^{-1}$ when $n, m \geq N_p$. Since $\{a_n\}$ is Cauchy-sequence, N_p is well defined. For all $p \in \mathbb{N}_{>0}$ we will define J_p as the interval

$$J_p := [a_{N_p} - (2p)^{-1}, a_{N_p} + (2p)^{-1}].$$

  Additionally we let $J_0 := J_1$. We now have a sequence of closed intervals $\{J_p\}$. However, it is unclear if it is also a nested sequence. To guarantee this property, we will define a new sequence $\{I_p\}$ where $I_0 := J_0$ and $I_p := J_0 \cap \ldots \cap J_p$ when $p \in \mathbb{N}_{>0}$. Clearly this sequence is nested. However, we still need to show that I_p is non-empty for each $p \in \mathbb{N}$.

  We will do so by showing that $a_{N_p} \in I_p$ for all $p \in \mathbb{N}$. Since N_p was minimal in its property, we have that $N_q \leq N_p$ when $q \leq p$. Therefore $a_{N_p} \in J_q$ whenever $q \leq p$. And since $I_p = J_0 \cap \ldots \cap J_p$

we get $a_{N_p} \in I_p$.

Finally we have $l(I_p) \leq l(J_p) = \frac{1}{p}$ for all $p \in \mathbb{N}_{>0}$. By Lemma A.1.9 we have that $l(I_p) \to 0$, since L is Archimedean. So $\{I_p\}$ is a nested sequence of closed intervals, whose length converges to 0. By Lemma A.1.13 we have the existence of some $a \in L$ such that $\cap I_p = \{a\}$. Since the interval I_p contains a_{N_p} and a for all $p \in \mathbb{N}$, and its length goes to zero for sufficiently large p, we have that $\{a_{N_p}\}_{p\geq 0}$ converges to a. So the Cauchy-sequence $\{a_n\}$ has a convergent subsequence $\{a_{N_p}\}_{p\geq 0}$. It is therefore itself convergent to the same limit $a \in L$.

– [$(iii) \implies (i)$] Let L be an Archimedean Cauchy-complete ordered field. We need to show that every non-empty subset of L that has an upper bound also as a least upper bound. Let A be such a subset of L, and let $R \in L$ be such an upper bound. We will construct a sequence of intervals $\{I_n\}$ with $I_n = [a_n, b_n]$ inductively such that b_n is an upper bound of A while a_n is not. Additionally we will have the length of each interval $l(I_n)$ going to zero for sufficiently large n.

Since A is non-empty, we can take an element $a \in A$. Let $I_0 := [a, R]$. Suppose that for some $n \in \mathbb{N}_{>0}$ the intervals $I_p = [a_p, b_p]$ for $p \leq n$ exist such that that b_p is an upper bound of A, while a_p is not, and $l(I_p) = 2^{-p} \cdot l(I_0)$. Then we will define the interval I_{n+1} as follows: let $c = a_n + 2^{-1} \cdot (b_n - a_n)$; if c is an upper bound of A, then let $I_{n+1} := [a_n, c]$. If it's not, then let $I_{n+1} := [c, b_n]$. In both cases $l(I_{n+1}) = 2^{-1} \cdot l(I_n) = 2^{-(n+1)} \cdot l(I_o)$.

To show that $l(I_n)$ does in fact converge to 0, note that $2^n = (1+1)^n \geq (1+n)$ for all $n \in \mathbb{N}$. Therefore, repeated application of Corollary A.1.2 gives us

$$\frac{1}{2^n} \leq \frac{1}{1+n} = \frac{1}{n} \cdot \frac{1}{1+\frac{1}{n}} \leq \frac{1}{n}.$$

We can make use of Lemma A.1.9 since L is Archimedean. Combined with the above inequality this lemma shows that $2^{-n} \to 0$ as $n \to \infty$. Therefore we have that $l(I_n) = 2^{-n} \cdot l(I_0)$ goes to 0 as well for sufficiently large n.

Since the length of the intervals goes to 0, we can look at the sequence $\{b_n\}$ and see that it is a Cauchy-sequence. Since L is Cauchy-complete there is some $b \in L$ such that $\{b_n\}$ converges to b. We can now show that this b has to be the least upper bound of A. If it were not, then either:

- b is not an upper bound at all. In this case there is a $a \in A$ such that $b < a$. But since b_n gets arbitrarily close to b, we get that $b_n < a$ for sufficiently large n. Since b_n is an upper bound by definition, we have a contradiction.
- b is not the *least* upper bound. Then there exist a c such that c is an upper bound and $c < b$. As the length of the intervals I_n goes to 0, there is an $N \in \mathbb{N}$ such that $b_N - a_N < b - c$. By property (O_1) of ordered fields we can rewrite this as $b_N < b + (a_N - c)$. Since a_n was by definition not an upper bound of A we have that $a_n < c$ for all n. Again using (O_1) this can be rewritten as $(a_n - c) < 0$. Combining these two results we get $b_N < b + (a_N - c) < b$. Since by definition b_n is a decreasing sequence we have that $b_n \leq b_N < b$. We see that for sufficiently large n there is a minimal distance of $b - B_N > 0$ between b_n and its limit. Therefore b is not a limit at all, contradicting its definition.

□

A well known result is that all Cauchy-complete Archimedean ordered fields are isomorphic to the real numbers. In this characterisation the real numbers are usually taken as the set of Dedekind-cuts in $\mathbb{Q}$ denoted by $\mathcal{D}$. With the above result in mind, this becomes the theorem:

Theorem A.1.14. *Every Dedekind-complete ordered field L is isomorphic to the real numbers.*

We won't proof the result here, we will just give an indication of how it's usually done. The first step is to extend the embedding of the natural numbers we saw in Lemma A.1.4 to an embedding of the rationals in all non-trivial ordered fields. Denote this embedding by $j : \mathbb{Q} \to L$. Then we define a function $f : L \to \mathcal{D}$ such that

$$f(x) := \{r \in \mathbb{Q} \mid j(r) < x\}.$$

Using the fact that L is Dedekind-complete we can show that $f(x)$ is indeed a Dedekind-cut and that f is a bijection.

Bibliography

[1] George Berkeley. *The Analyst: Or, a Discourse Adressed to an Infidel Mathmatician; Wherein it is Examined Whether the Object, Principles, and Inferences of the Modern Analysis are More Distinctly Concieved, or More Evidently Deduced, than Religious Mysteries and Points of Faith.* Consulted via “https://www.maths.tcd.ie/pub/HistMath/People/Berkeley/Analyst/Analyst.pdf”, on 14th June 2020. London, 1734.

[2] Lothar Sebastian Krapp. *Constructions of the Real Numbers: A Set Theoretical Approach.* University of Oxford. Consulted via “http://www.math.uni-konstanz.de/~krapp/Constructions_of_the_real_numbers.pdf”, on 3rd June 2020. Oxford, 2014.

[3] Tom Lindstrøm. *An Invitation to Nonstandard Analysis.* Accessed through “https://www.cambridge.org/core/books/nonstandard-analysis-and-its-applications/an-invitation-to-nonstandard-analysis/EE83E71E747D7B2DDFCBC07A286502FF”, on 14th June 2020. Cambridge University Press, 1988.

Index