University of Helsinki
Faculty of Science
Master's Programme in Mathematics and Statistics

Master's thesis

Introduction to Nonstandard Analysis and Loeb Measures

Oskar Björkman

Supervisor:
Hans-Olav Tylli

September 2025

Title: Introduction to Nonstandard Analysis and Loeb Measures
Author: Oskar Björkman
Month and year: September 2025
Page count: 74 pp.

Abstract:

The goal of this work is to provide an introduction to nonstandard analysis and Loeb measures, preparing the reader for more advanced studies in nonstandard analysis, and nonstandard probability theory in particular. To this end this work is divided into three parts.

In the first part we give a presentation of the elements of the nonstandard framework for analysis. We construct the nonstandard theory using the superstructure approach and prove a central result, the transfer principle, of the nonstandard approach to analysis. The presentation is simplified slightly by giving a direct proof, instead of one relying on Los's theorem as is usually done. The construction is followed by a discussion of the fundamental concepts and tools needed for nonstandard analysis. The first section concludes by developing the basics of nonstandard topology and nonstandard real analysis.

In the second part we give an account of the theory of unbounded Loeb measures. Because many applications of Loeb measures are in probability theory there are few published accounts of the full theory of unbounded Loeb measures. We begin by detailing the underlying construction. After the construction of Loeb measures we construct a nonstandard Lebesgue measure using the Loeb measure machinery. Once this is done the rest of this part is spent presenting the elements of the theory of Loeb measurable functions and the related integration theory.

In the third and final part we apply the theory developed so far to give a presentation of the nonstandard approach to the theory of one dimensional Brownian motion and the associated Wiener measure.

Finally, while we have strived to give an elementary account of the subjects presented in this work, certain prerequisites nevertheless exist. Notably, the introduction to nonstandard analysis requires a certain facility with set theory and the material on Loeb measures makes use of some classical results from measure theory without proving them. Despite this, the vast majority of the theory of Loeb measures can be read without any prior familiarity with measure theory. The final part on Brownian motion requires an understanding of Loeb measures and nonstandard analysis.

Keywords: nonstandard analysis, loeb measure, brownian motion

Contents

Introduction

Historically, infinitesimals date back to Archimedes, however, they played their most pivotal role in the development of calculus. As calculus matured, the need to provide the subject a rigorous foundation grew. Ultimately the limit based approach developed by Cauchy, Weierstraß, etc. cemented itself as the preferred formalism for analysis. Consequently infinitesimal arguments fell out of favour and a general sense that the notion of infinitesimals might contain some contradiction settled in. Despite this their standing as heuristic tools was not terribly diminished.

The infinitesimals were finally vindicated by Abraham Robinson in the 1960s. When he discovered a productive framework for working with infinitesimals. His framework gave rise to the branch of mathematical analysis now known as nonstandard[1]. About a decade later another approach to analysis that features infinitesimals was also developed. Namely, synthetic differential geometry.

In nonstandard analysis we are primarily concerned with the sort of infinitesimals and infinitesimal arguments that are, perhaps primarily, due to Leibniz. For us infinitesimals and, their reciprocals, the infinites are best thought of as idealised entities. That is they are useful mathematical fictions. Much like the transcendental real numbers, the complex numbers, etc. Just as you cannot write out a numerical representation of π the infinitesimal numbers do not admit a numerical representation in the usual sense. In particular, the idea is that the usual rules extend to these new elements and vice versa.

Naturally, nonstandard analysis features a construction of a non-archimedean number system. More importantly it places them in a context where the idea that they have all the properties of the regular numbers is established rigorously. This correspondence between what is true for regular numbers and the extended idealised numbers is called the *transfer principle*. In fact, nonstandard analysis and the transfer principle give us, as we shall see, much more. The main idea is to construct an extended mathematical universe where all the new idealised elements are present which is connected to the original universe by the transfer principle.

Despite the undeniable intuitive appeal of infinitesimals, nonstandard analysis presents a pedagogical challenge. Its foundations consist primarily of logic, set theory, and model theory – branches of mathematics that are seldom used in standard mathematical analysis.

In spite of this obstacle nonstandard analysis has, perhaps due to the intuitive nature of the arguments it provides a formal language for, made and continues to make contributions to many branches of mathematics Among others it has made contributions to: potential theory, number theory, mathematical physics, and mathematical economics and finance. For examples of these see *Nonstandard Analysis for the Working Mathematician* (Loeb, Wolff, et al. 2000).

The first part of this document is dedicated to an elementary treatment of the foundations of nonstandard analysis. Here we construct the nonstandard universe, prove that it satisfies the transfer principle, and introduce the basic tools for making "nonstandard" arguments.

In the second part, we apply the tools we developed in the first part to provide an elementary treatment of Loeb measures. Loeb measures have found wide-ranging applications in probability

[1] The name nonstandard analysis is a bit unfortunate and thus in good company together with complex analysis, imaginary numbers, and irrational numbers.

theory, particularly in stochastic analysis. Their applications constitute, arguably, most of the contributions made by nonstandard analysis.

Finally, following Anderson 1976, we apply the nonstandard framework and Loeb measures to the problem of modelling Brownian motion.

While we treat what constitutes the necessary foundations, we stop just short of some of the more interesting applications of nonstandard analysis. For instance, the Ito integral can be given a very elegant and satisfying characterisation using nonstandard analysis. For such an account, see for instance Anderson 1976 or Albeverio et al. 1986. Either should be approachable for a reader familiar with the material we present.

Chapter 1

Nonstandard Analysis

In this part we aim to introduce the subject of nonstandard analysis as succinctly as possible, while remaining rigorous. In our opinion trying to explain the fundamental construction or the the fundamental result (i.e. the transfer principle) informally is unfruitful. Instead we present the core machinery as succinctly as possible before explaining how it is to be employed. Once the core of the formalism is understood the subject starts to open up properly. We must warn the reader that initially there is a certain lack of context and a fairly brisque pace. Once we get past the transfer principle things slow down a bit and we start building context.

In 1.1 The Construction we detail the superstructure approach to constructing the nonstandard "universe" and prove the transfer principle for it.[1] The superstructure approach has become the de facto standard for introductions to nonstandard analysis. At least, when their target audience is the mathematical analysts. Robinson's original approach based on type theory, etc. is a bit more general, but requires far more mathematics generally unfamiliar to those specialised in mathematical analysis.

The subsequent sections address elementary nonstandard terminology and theory, along with the basic notions of nonstandard analysis. Here we attempt to cram in all the consequences of the transfer principle, as well as some additional tools, that are needed in practice when making nonstandard arguments.

By the end of this the reader should, in theory, have a basic level of fluency in the nonstandard framework.

[1]The transfer principle is most frequently proven as a corollary to Łoś's theorem, particularly when the superstructure construction is employed. Here, we present a direct proof.

1.1 The Construction

In the construction the set X will stand for the set of atomic entities we want to reason about using nonstandard arguments. For instance, if we want elementary nonstandard real analysis, i.e. statements about functions, relations, etc. on the real numbers, we would choose $X = \mathbb{R}$. Importantly, we will not be able to make nonstandard arguments about the internals of the elements of X. That is if we choose $X = \mathbb{R}$ as a set of Dedekind cuts we will not be able to make nonstandard arguments that rely on the fact that the elements of $\mathbb{R}$ are sets. (In the case of Dedekind cuts they would be subsets of $\mathbb{Q}$. Naturally, should we need to make such arguments we can just choose $X = \mathbb{Q}$ instead.) In practice the choice of X is not of any real concern. We will simply make our arguments in the knowledge that an appropriate choice of X could be made and determined from the arguments we made a posteriori, should the need or desire arise.

The objects of the first order theory on the set X consists of the relations and functions on this set. We therefore want to include all these objects in our construction. The set of all (finitary) relations, etc. on X is called the *superstructure* on X.

Definition 1.1.1 (Superstructure)**.** Suppose X is a non-empty set. Then the superstructure on X is defined

$$V(X) := \bigcup_{n}^{\mathbb{N}} V_n(X),$$

where

$$V_0(X) = X,\ V_{n+1}(X) = V_n(X) \cup \mathcal{P}(V_n(X)).$$

At times the convention that $V_{-1}(X) = \emptyset$ will be used. Finally, for each $x \in V(X)$ the grade $\mathfrak{g}(x)$ of x is the lowest index n for which $x \in V_n(X)$.

As a matter of course we shall assume that $\mathbb{N} \subset X$. Due to how frequently the natural numbers are employed.

In general we will conceive of the natural numbers in the style of von Neumann by the recursive scheme: $0 = \emptyset$, $n = \{k \mid k < n\}$. This shows up primarily in our notation. In particular, the following expressions are equivalent $\bigcup_m^n A_m = \bigcup_{m \in n} A_m$, where A is some n-tuple of sets and $n \in \mathbb{N}$. The particular encoding of $\mathbb{N}$ you would include in X for a particular application will vary depending on the sorts of arguments you need to make. Thus when we say that $\mathbb{N} \subset X$ we are saying that we assume that you have some appropriate notion of the natural numbers in X.

The functions, relations, and tuples are included in the superstructure on X in the usual set theoretical sense. Tuples are encoded:

Definition 1.1.2 (n-tuple)**.** A 0-tuple is just the empty set. The $n+1$-tuple $\langle a_i \rangle_i^{n+1}$ is then defined recursively by

$$\langle a_i \rangle_i^{n+1} = \{\{a_n\} \cup \langle a_i \rangle_i^n, \} \cup \langle a_i \rangle_i^n.$$

For instance,

$$\begin{aligned} \langle a, b, c \rangle &= \{\{c\} \cup \langle a, b \rangle\} \cup \langle a, b \rangle \\ &= \{\{c\} \cup \{b\} \cup \langle a \rangle\} \cup \{\{b\} \cup \langle a \rangle\} \cup \langle a \rangle \\ &= \{\{c\} \cup \{b\} \cup \{a\} \cup \emptyset\} \cup \{\{b\} \cup \{a\} \cup \emptyset\} \cup \{a\} \\ &= \{\{c, b, a\}, \{b, a\}, a\}. \end{aligned}$$

Moreover, relations are encoded:

Definition 1.1.3 (Relations)**.** The *cartesian product* of $A_i, i \in n$ is defined as the set of tuples $\langle a_i \rangle, i \in n$ such that $a_i \in A_i$ for each $i \in n$ and denoted

$$\prod_i^n A_i.$$

A *n-ary relation* on $(A_i)_i^n, i \in n$ is any subset of $\prod_i^n A_n$.

Before moving on we note that;

Proposition 1.1.4. *For any X the set of finite subsets of X is a subset of the superstructure $V(X)$. In brief $\mathcal{P}_F(V(X)) \subset V(X)$.*

Proof. For any finite subset $U \subseteq V(X)$ there is an element with the highest grade. Suppose this maximal grade is $n \in \mathbb{N}$. Then $U \in \mathcal{P}(V_n(X)) \subseteq V_{n+1}(X)$. The proposition follows. □

The key concept in the construction will be that of a filter, ultrafilter in particular.

Definition 1.1.5 (filter & ultrafilter)**.** A subset $\mathcal{U} \subset \mathcal{P}(\mathcal{I})$ is called a *filter* on the set $\mathcal{I}$ if

$$\mathcal{I} \in \mathcal{U}, \emptyset \notin \mathcal{U}, \tag{1.1}$$

$$\text{if } A \in \mathcal{U} \text{ and } A \subset B \subset \mathcal{I}, \text{ then } B \in \mathcal{U}, \qquad \text{(closed under supersets)}$$

$$\text{if } A, B \in \mathcal{U}, \text{ then } A \cap B \in \mathcal{U}. \qquad \text{(closed under intersection)}$$

A filter is called an *ultrafilter* if it also satisfies

$$\text{if } A \subseteq \mathcal{I} \text{ ,then } A \in \mathcal{U} \text{ or } A^{\complement} = \mathcal{I} \setminus A \in \mathcal{U}. \tag{1.2}$$

Finally, a filter is *cofinite* if it includes all subsets of $\mathcal{I}$ which have finite complement.

In the context of a filter $\mathcal{U}$ on an index set $\mathcal{I}$ the complement of a subset $u \subset \mathcal{I}$ with respect to the index set will be denoted $u^{\complement}$, i.e. $u^{\complement} = \mathcal{I} \setminus u$. We will not need much of the theory of filters but the following two propositions will be employed repeatedly.

Proposition 1.1.6. *No filter can contain a set and its complement.*

Proof. Let $\mathcal{U}$ be a filter and $u \in \mathcal{U}$ then if $u^{\complement} \in \mathcal{U}$ it follows that $\emptyset = u \cap u^{\complement} \in \mathcal{U}$. But, then $\mathcal{U}$ isn't a filter since filters do not contain the empty set. □

Proposition 1.1.7 (ultrafilter exclusion)**.** *Let $\mathcal{U}$ be an ultrafilter on $\mathcal{I}$ and A be a set of $n \in \mathbb{N}$ subsets of $\mathcal{I}$. Then, the union, $\bigcup_a^A a$, is in the ultrafilter $\mathcal{U}$ if and only if some element $a \in A$ is in the ultrafilter $\mathcal{U}$. Moreover, if the elements of A are disjoint, then their union, $\bigcup_a^A u$, is in the ultrafilter $\mathcal{U}$ if and only if exactly one of the elements $a \in A$ is in the ultrafilter $\mathcal{U}$.*

Proof. Let $\mathcal{U}$, $\mathcal{I}$, and A be as in the proposition statement.

We begin by showing that at least one of the elements of $a \in A$ is in the ultrafilter $\mathcal{U}$ if their union $\bigcup_a^A a$ is in $\mathcal{U}$. Suppose, that no element $a \in A$ is in $\mathcal{U}$. Then, the complement $a^{\complement}$ is in the ultrafilter for each $a \in A$, since $\mathcal{U}$ is an ultrafilter. Moreover, the intersection $\bigcap_a^A a^{\complement}$ of the complements must be in the ultrafilter, since filters are closed under finite intersection. However, now it follows that

$$\bigcup_a^A a = \left(\bigcap_a^A a^{\complement} \right)^{\complement} \notin \mathcal{U},$$

since $\mathcal{U}$ is an ultrafilter. This contradicts the assumption that the union is in the ultrafilter. Thus, at least some elment $a \in A$ must be in the ultrafilter $\mathcal{U}$ if the union $\bigcup_a^A a$ is in the ultrafilter $\mathcal{U}$.

Now, suppose that the elements of A are also disjoint. It follows that $a^{\complement} \supseteq a'$ for all $a, a' \in A$ whenever $a \neq a'$. Let $a' \in A$ be such that a' is in the ultrafilter, such an element exists by the first part of this proof. Then, for each $a \in A \setminus \{a'\}$ the set $a^{\complement}$ is in the ultrafilter, since filters are closed under supersets. Moreover, for each $a \in A \setminus \{a'\}$, a cannot be in the ultrafilter, by Proposition 1.1.6. In other words, exactly one element of A is in the ultrafilter $\mathcal{U}$.

Conversely, if some element $a \in A$ is in the ultrafilter $\mathcal{U}$, then $\bigcup_a^A a \supseteq a$ and so $\bigcup_a^A a \in \mathcal{U}$, since $\bigcup_a^A a \subseteq \mathcal{I}$ and filters are closed under supersets. □

The following will turn out to be the central structure of the construction.

Definition 1.1.8 (Ultrapower)**.** The ultrapower of X with respect to the ultrafilter $\mathcal{U}$ on the index set $\mathcal{I}$ is defined as the set of equivalence classes

$$[a]_{\mathcal{U}} := \{a' \in X^{\mathcal{I}} \mid \{i \in \mathcal{I} \mid a(i) = a'(i)\} \in \mathcal{U}\}, \quad a \in X^{\mathcal{I}}.$$

We denote the ultrapower by

$$\prod\nolimits_{\mathcal{U}} X = \{[a]_{\mathcal{U}} \mid a \in X^{\mathcal{I}}\}.$$

Before we proceed we make sure that the definition makes sense.

Proposition 1.1.9. *The ultrapower is well-defined.*

Proof. We need to show that the "equivalence classes" of our definition are in fact equivalence classes. That is

$$a =_{\mathcal{U}} a' := \{i \in \mathcal{I} \mid a(i) = a'(i)\} \in \mathcal{U}$$

forms an equivalence relation. It is clearly symmetric and reflexive by definition. It remains only to be shown that it is transitive. Let $a =_{\mathcal{U}} b$ and $b =_{\mathcal{U}} c$ meaning the sets

$$\{i \in \mathcal{I} \mid a(i) = b(i)\}, \{i \in \mathcal{I} \mid b(i) = c(i)\} \in \mathcal{U},$$

by definition. It follows that their intersection is in $\mathcal{U}$ as well. Now since

$$\{i \in \mathcal{I} \mid a(i) = c(i)\} \supset \{i \in \mathcal{I} \mid a(i) = b(i)\} \cap \{i \in \mathcal{I} \mid b(i) = c(i)\}.$$

It follows that the lefthand side is in $\mathcal{U}$, since filters are closed under supersets. Which is equivalent to $a =_{\mathcal{U}} c$ by definition. □

Let us briefly consider the role that these objects will play. We want to construct some object with "all the limits". Heuristically, we take all the "sequences", i.e. $X^{\mathcal{I}}$. And then identify the ones that are equal "almost everywhere", i.e. we take the quotient with respect to a cofinite ultrafilter. For instance, we could take $\mathbb{N}$ as our index set $\mathcal{I}$ and an ultrafilter which identifies all sequences that only differ at a finite set of indices, i.e. a cofinite ultrafilter. This would already give us a model with some of the properties we want. For instance if $X = \mathbb{R}$ we get a model of the hyperreals. (However, it turns out $\mathbb{N}$ is too small for the general context. More on this in section 1.4 Saturation Principles.)

Now, there is a natural way to embed X in $\prod_{\mathcal{U}} X$ by taking each element to the equivalence class of its constant function, it is not so clear how one would extend this to the superstructures in general. To this end we define;

Definition 1.1.10 (Bounded Ultrapower)**.** The bounded ultrapower on X over the filter $\mathcal{U}$ on the index set $\mathcal{I}$ is defined

$$\overline{\prod}_{\mathcal{U}} X := \bigcup_{n}^{\mathbb{N}} \prod\nolimits_{\mathcal{U}} (V_n(X) \setminus V_{n-1}(X)). \tag{1.3}$$

The bounded ultrapower does not result in a superstructure. We will choose to remedy this in time. (See *Introduction to the Theory of Infinitesimals* (Stroyan and Luxemburg 1976) for an approach which works directly on the bounded ultrapower.) A first step in recovering that structure is the following relation.

Definition 1.1.11 ($\in_{\mathcal{U}}$)**.** Let $a, b \in \overline{\prod}_{\mathcal{U}} X$, $a = [x]_{\mathcal{U}}$, and $b = [y]_{\mathcal{U}}$ then we say a is a $\mathcal{U}$-element of b, denoted $a \in_{\mathcal{U}} b$, if $\{i \in \mathcal{I} \mid x(i) \in y(i)\} \in \mathcal{U}$.

Once again we make sure that the notion is well-defined.

Proposition 1.1.12. *The relation $\in_{\mathcal{U}}$ is well-defined, meaning if $[x]_{\mathcal{U}} \in_{\mathcal{U}} [y]_{\mathcal{U}}$, $[x']_{\mathcal{U}} = [x]_{\mathcal{U}}$, and $[y']_{\mathcal{U}} = [y]_{\mathcal{U}}$ then $[x']_{\mathcal{U}} \in_{\mathcal{U}} [y']_{\mathcal{U}}$.*

Proof. By definition the sets $A = \{i \in \mathcal{I} \mid x(i) = x'(i)\}$, $B = \{i \in \mathcal{I} \mid y(i) = y'(i)$, and $C = \{i \in \mathcal{I} \mid x(i) \in y(i)\}$ are in the ultrafilter $\mathcal{U}$. Since filters are closed under finite intersection so is their intersection $A \cap B \cap C$. Finally since

$$\{i \in \mathcal{I} \mid x'(i) \in y'(i)\} \supset A \cap B \cap C$$

and filters are closed under supersets it follows that the set on the lefthand side is in $\mathcal{U}$. Thus $[x']_{\mathcal{U}} \in_{\mathcal{U}} [y']_{\mathcal{U}}$, by definition. □

An immediate but important property of the $\mathcal{U}$-membership relation, $\in_{\mathcal{U}}$, is that it preserves the hierarchy in the following sense;

Proposition 1.1.13. *If $a, b \in \overline{\prod}_{\mathcal{U}} X$, $a \in_{\mathcal{U}} b$, and $b \in \prod_{\mathcal{U}}(V_n(X) \setminus V_{n-1}(X))$, then $a \in \prod_{\mathcal{U}}(V_k(X) \setminus V_{k-1}(X))$ for some $k < n$.*

Proof. Let $[x]_{\mathcal{U}}, [y]_{\mathcal{U}} \in \overline{\prod}_{\mathcal{U}} X$. If $[x]_{\mathcal{U}} \in_{\mathcal{U}} [y]_{\mathcal{U}}$ then $\{i \in \mathcal{I} \mid x(i) \in y(i)\} \in \mathcal{U}$, by definition. Now, if $[y]_{\mathcal{U}} \in \prod_{\mathcal{U}}(V_n(X) \setminus V_{n-1}(X))$, $n \in \mathbb{N}$ then each $y(i) \in (V_n(X) \setminus V_{n-1}(X))$. Thus, since $x(i) \in y(i)$ for each $x(i)$ it follows that each $x(i)$ is in $V_{n-1}(X)$. Finally, since $[x]_{\mathcal{U}} \in \prod_{\mathcal{U}}(V_k(X) \setminus V_{k-1}(X))$ for some k, by definition, it follows that $k < n$. □

This suggests the following definition;

Definition 1.1.14 ($\mathcal{U}$-set)**.** The elements of $\overline{\prod}_{\mathcal{U}} X \setminus \prod_{\mathcal{U}} X$ are called $\mathcal{U}$-sets. A $\mathcal{U}$-set is called finite if it only has a finite number of $\mathcal{U}$-elements.

Note that we exclude the elements of $\prod_{\mathcal{U}} X$ from the $\mathcal{U}$-sets because they cannot contain any $\mathcal{U}$-elements, by Proposition 1.1.13. It turns out that the elements of the bounded ultrapower which are $\mathcal{U}$-sets are uniquely defined by their $\mathcal{U}$-elements.

Proposition 1.1.15 ($\mathcal{U}$-set representation)**.** *Each $\mathcal{U}$-set is uniquely defined by the set of its $\mathcal{U}$-elements. Therefore if $\{a_j \mid j \in \mathcal{J}\}$ is the set of $\mathcal{U}$-elements of some $\mathcal{U}$-set, then we can unambiguously refer to that $\mathcal{U}$-set by the representation*

$$\{a_j \mid j \in \mathcal{J}\}_{\mathcal{U}}.$$

Proof. Let $[b]_{\mathcal{U}}$ and $[b']_{\mathcal{U}}$ be $\mathcal{U}$-sets such that $[b]_{\mathcal{U}} \neq [b']_{\mathcal{U}}$. Then, $u = \{i \in \mathcal{I} \mid b(i) = b'(i)\} \notin \mathcal{U}$, by definition, and therefore $\{i \in \mathcal{I} \mid b(i) \neq b'(i)\} \in \mathcal{U}$, since $\mathcal{U}$ is an ultrafilter. Now,

$$\begin{aligned} u &= \{i \in \mathcal{I} \mid (b(i) \setminus b'(i)) \cup (b'(i) \setminus b(i)) \neq \emptyset\} \\ &= \{i \in \mathcal{I} \mid b(i) \setminus b'(i) \neq \emptyset\} \cup \{i \in \mathcal{I} \mid b'(i) \setminus b(i) \neq \emptyset\}, \end{aligned}$$

and so one of the two sets in the final righthand union must be in $\mathcal{U}$, by ultrafilter exclusion (1.1.7). Suppose, without loss of generality, that

$$u' = \{i \in \mathcal{I} \mid b(i) \setminus b'(i) \neq \emptyset\} \in \mathcal{U}.$$

Now

$$u' = \bigcup_{k}^{m} u'_k = \bigcup_{k}^{m} \{i \in u' \mid b(i) \in (V_k(X) \setminus V_{k-1}(X))\},$$

and so one of the $u'_k \in \mathcal{U}$, again by ultrafilter exclusion (1.1.7). Assuming $u'_h \in \mathcal{U}$ define

$$v(i) = \begin{cases} c(i) \in b(i) \setminus b'(i), & i \in u'_h, \\ c(i) \in V_k(X) \setminus V_{k-1}(X), & i \notin u'_h. \end{cases}$$

Then $[c]_\mathcal{U} \in \prod_\mathcal{U}(V_k(X) \setminus V_{k-1}(X))$ and $[c]_\mathcal{U} \in_\mathcal{U} [b]_\mathcal{U}$, by construction. We now complete the argument by showing that $[c]_\mathcal{U} \notin [b']_\mathcal{U}$. But, this is equivalent to $\{i \in \mathcal{I} \mid c(i) \notin b'(i)\} \in \mathcal{U}$, since $\mathcal{U}$ is an ultrafilter. Which follows from the observation that $\{i \in \mathcal{I} \mid c(i) \notin b'(i)\} \supset u'_h$. We have now shown that if $[b]_\mathcal{U} \neq [b']_\mathcal{U}$ then one of the two has a $\mathcal{U}$-element which is not in the other. And so the $\mathcal{U}$-set representation is unique for elements of $\overline{\prod}_\mathcal{U} X \setminus \prod_\mathcal{U} X$. □

Note that not every set of "$\mathcal{U}$-elements" is necessarily possible. In particular,

$$\{a_i \mid i \in \mathcal{J}\}_\mathcal{U},$$

doesn't make sense if no $\mathcal{U}$-set with exactly the $\mathcal{U}$-elements $\{a_i \mid i \in \mathcal{J}\}$ exists. Thus, whenever we use the notation the existence of such a $\mathcal{U}$-set is implicitly assumed.

Next we turn our attention to the natural embedding.

Definition 1.1.16. The natural embedding $\iota : V(x) \to \overline{\prod}_\mathcal{U} X$ of the superstructure $V(X)$ into the bounded ultrapower $\overline{\prod}_\mathcal{U} X$ is defined as the function which takes each element $a \in V_n(X) \setminus V_{n-1}(X)$ to the the equivalence class $[\hat{a}]_\mathcal{U} \in \prod_\mathcal{U}(V_n(X) \setminus V_{n-1}(X))$, of the constant function $\hat{a} : \mathcal{I} \to V_n(X); i \mapsto a$.

Which has the following properties;

Proposition 1.1.17. *The embedding ι has the properties:*

$$a \in b \Leftrightarrow \iota(a) \in_\mathcal{U} \iota(b) \quad \forall a, b \in V(X), \qquad \text{(monomorphism of membership)}$$
$$a = b \Leftrightarrow \iota(a) = \iota(b) \quad \forall a, b \in V(X), \qquad \text{(monomorphism of equality)}$$
$$\iota(\{a_i \in V(X) \mid i \in n\}) = \{\iota(a_i) \mid i \in n\}_\mathcal{U}. \qquad \text{(monomorphism of finite sets)}$$

Throughout this proof we will use the $\hat{a}$ notatin for the constant function of a set $a \in V(X)$ as defined in the definition of the natural embedding.

Proof. Let $a, b \in V(X)$ then $a, b \in (V_n(X) - V_{n-1}(X))$ for some $n \in \mathbb{N}$. Now

$$\begin{aligned}\iota(a) \in_\mathcal{U} \iota(b) &\Leftrightarrow \{i \in \mathcal{I} \mid \hat{a}(i) \in \hat{b}(i)\} \in \mathcal{U}\\ &\Leftrightarrow \{i \in \mathcal{I} \mid a \in b\} \in \mathcal{U}\\ &\Leftrightarrow a \in b.\end{aligned}$$

The second equivalence is obtained in the same manner. For the third, let $a = \{a_i \in V(X) \mid i \in n\} \in \mathcal{P}_F(V(X))$ and $[x]_\mathcal{U} \in \overline{\prod}_\mathcal{U} X$ then

$$\begin{aligned}&[x]_\mathcal{U} \in_\mathcal{U} \iota(\{a_j \in V(X) \mid j \in n\}) \Leftrightarrow \{i \in \mathcal{I} \mid x(i) \in \hat{a}(i)\} \in \mathcal{U}\\ &\quad\Leftrightarrow \{i \in \mathcal{I} \mid x(i) \in a\} \in \mathcal{U}\\ &\quad\Leftrightarrow \{i \in \mathcal{I} \mid \exists j \in n(x(i) = a_j)\} \in \mathcal{U}\\ &\quad\Leftrightarrow \bigcup_j^n \{i \in \mathcal{I} \mid x(i) = a_j\} \in \mathcal{U}\\ &\quad\Leftrightarrow \exists! j \in n(\{i \in \mathcal{I} \mid x(i) = a_j\} \in \mathcal{U}), \quad \text{by ultrafilter exclusion (1.1.7)},\\ &\quad\Leftrightarrow \exists! j \in n([x]_\mathcal{U} \in [\widehat{a_j}]_\mathcal{U})\\ &\quad\Leftrightarrow \exists! j \in n([x]_\mathcal{U} = \iota(a_j))\\ &\quad\Leftrightarrow [x]_\mathcal{U} \in \iota[a].\end{aligned}$$

□

Observe that the content of being a monomorphism of equality is really just to say that ι is an injection. The $\Rightarrow$ part of it is true for any function. And the $\Leftarrow$ part is the defining property of an injection on its domain, in this case $V(X)$.

The following gives us the correct map from the bounded ultrapower to the superstructure of the ultrapower; The idea is just to turn $\in_{\mathcal{U}}$ back into the regular notion of membership $\in$. The function is an instance of a Mostowski collapse function. Since it is the only such function we shall need we will refer to it as the Mostowski collapse function.

Definition 1.1.18 (Mostowski collapse)**.** The Mostowski collapse function $M : \overline{\prod}_{\mathcal{U}} X \to V(\prod_{\mathcal{U}} X)$ acts as follows;

$$M([b]_{\mathcal{U}}) = \begin{cases} [b]_{\mathcal{U}}, & [b]_{\mathcal{U}} \in \prod_{\mathcal{U}} X, \\ \{M([a]_{\mathcal{U}}) \mid [a]_{\mathcal{U}} \in_{\mathcal{U}} [b]_{\mathcal{U}}\}, & [b]_{\mathcal{U}} \in \overline{\prod}_{\mathcal{U}} X \setminus \prod_{\mathcal{U}} X. \end{cases}$$

This function turns out to have the same important monomorphism properties that we proved for ι;

Proposition 1.1.19. *The Mostowski collapse function has the properties:*

$$a \in_{\mathcal{U}} b \Leftrightarrow M(a) \in M(b), \quad \forall a, b \in \overline{\prod}_{\mathcal{U}} X, \qquad \text{(monomorphism of membership)}$$

$$a = b \Leftrightarrow M(a) = M(b), \quad \forall a, b \in \overline{\prod}_{\mathcal{U}} X, \qquad \text{(monomorphism of equality)}$$

$$M(\{a_i \in \overline{\prod}_{\mathcal{U}} \mid i \in n\}_{\mathcal{U}}) = \{M(a_i) \mid i \in n\}, \quad \forall n \in \mathbb{N} \qquad \text{(monomorphism of finite sets)}$$

Proof. Let $[a]_{\mathcal{U}}, [b]_{\mathcal{U}} \in \overline{\prod}_{\mathcal{U}} X$.

We begin by proving the first equivalence. First note that

$$M([a]_{\mathcal{U}}) \in M([b]_{\mathcal{U}}) \Rightarrow [a]_{\mathcal{U}} \in_{\mathcal{U}} [b]_{\mathcal{U}},$$

by definition. Conversely, suppose $[a]_{\mathcal{U}} \in [b]_{\mathcal{U}}$. Observe that $[b]_{\mathcal{U}} \in \prod_{\mathcal{U}}(V_n(X) \setminus V_{n-1}(X))$ for some $n \in \mathbb{N}$, by definition. If $n = 0$ then $[a]_{\mathcal{U}} \in \prod_{\mathcal{U}}(V_k(X) \setminus V_{k-1}(X))$ for some $k < n$, by Proposition 1.1.13. But, no such k exists. So the proposition holds. If $n > 0$. Then $[a]_{\mathcal{U}} \in \prod_{\mathcal{U}}(V_k(X) \setminus V_{k-1}(X))$ for some $k < n$, again by Proposition 1.1.13, and so $M([a]_{\mathcal{U}}) \in M([b]_{\mathcal{U}})$, by definition. This proves the first equivalence.

We now prove the second equivalence. Only the converse ($\Leftarrow$) is nontrivial. Let $[a]_{\mathcal{U}}, [b]_{\mathcal{U}} \in \bigcup_k^m \prod_{\mathcal{U}} V_k(X)$. The proof is by induction on m. Let $m = 0$. Then the proposition is trivially true since on $\bigcup_k^m \prod_{\mathcal{U}}(V_k(X) \setminus V_{k-1}(X)) = \prod_{\mathcal{U}} X$ the function M acts as the identity function.

Suppose then that the proposition holds up to $m = n$. Now suppose $[a]_{\mathcal{U}}, [b]_{\mathcal{U}} \in \bigcup_k^{n+1} \prod_{\mathcal{U}}(V_k(X) \setminus V_{k-1}(X))$ and $[a]_{\mathcal{U}} \neq [b]_{\mathcal{U}}$. If both are in $\prod_{\mathcal{U}} X$ then this is covered by the base step. If only one is in $\prod_{\mathcal{U}} X$ the property then $M([a]_{\mathcal{U}}) \neq M([b]_{\mathcal{U}})$, by the definition of M. Now if $[a]_{\mathcal{U}}, [b]_{\mathcal{U}} \in \bigcup_k^{n+1} \prod_{\mathcal{U}} X \setminus \prod_{\mathcal{U}} X$ Then either one, let it be $[a]_{\mathcal{U}}$, has a $\mathcal{U}$-element $[c]_{\mathcal{U}}$ not contained in the other, by $\mathcal{U}$-set representation (1.1.15). Now $M([c]_{\mathcal{U}}) \in M([a]_{\mathcal{U}})$, by definition. Moreover $M([c]_{\mathcal{U}}) \notin M([b]_{\mathcal{U}})$, by the induction hypothesis since $[c]_{\mathcal{U}} \in \bigcup_k^{n+1} \prod_{\mathcal{U}} X$. This proves the second equivalence.

For the third equivalence, let $[b]_{\mathcal{U}}$ be the finite $\mathcal{U}$-set $\{[a_i]_{\mathcal{U}} \mid i \in n\}_{\mathcal{U}}$, $n \in \mathbb{N}$. Then

$$\begin{aligned} M([b]_{\mathcal{U}}) &= \{M([x]_{\mathcal{U}}) \mid [x]_{\mathcal{U}} \in_{\mathcal{U}} [b]_{\mathcal{U}}\} \\ &= \{M([x]_{\mathcal{U}}) \mid [x]_{\mathcal{U}} \in \{[a_i]_{\mathcal{U}} \mid i \in n\})\} \\ &= \{M([a_i]_{\mathcal{U}}) \mid i \in n\}. \end{aligned}$$

□

From this point on, let ${}^*X = \prod_{\mathcal{U}} X$. Finally we can define the transfer function which will take us from our standard model to the nonstandard one.

Definition 1.1.20 (transfer function)**.** The *transfer function*, ${}^* : V(X) \to V({}^*X)$, is defined as the composition of the embedding ι and the Mostowski collapse function M,

$$ {}^* = M \circ \iota. $$

Finally, we prove that the transfer function also satisfies the monomorphism properties we proved for the embedding and the Mostowski collapse function.

Proposition 1.1.21.

$$\begin{aligned} a \in b \Leftrightarrow {}^*a \in {}^*b, \quad \forall a, b \in V(X) & \qquad \text{(monomorphism of membership)} \\ a = b \Leftrightarrow {}^*a = {}^*b, \quad \forall a, b \in V(x) & \qquad \text{(monomorphism of equality)} \\ {}^*\{a_i \in V(X) \mid i \in n\} = \{{}^*a_i \in V(X) \mid i \in n\}. & \qquad \text{(monomorphism of finite sets)} \end{aligned}$$

Proof. The properties follow from the fact that both ι and M have the properties in question and the properties are inherited by the composite. Let $a, b \in V(X)$ and $c \in \mathcal{P}_F(V(X))$. Now by direct application of the monomorphism of membership, quality, and finite sets properties proved in Proposition 1.1.17 and Proposition 1.1.19 we get

$$\begin{aligned} a \in b \Leftrightarrow \iota(a) \in_{\mathcal{U}} \iota(b) \Leftrightarrow M(\iota(a)) \in M(\iota(b)) \Leftrightarrow {}^*a \in {}^*b, \\ a = b \Leftrightarrow \iota(a) = \iota(b) \Leftrightarrow M(\iota(a)) = M(\iota(b)) \Leftrightarrow {}^*a = {}^*b, \end{aligned}$$

and

$$\begin{aligned} {}^*\{a_i \in V(X) \mid i \in n\} &= M(\iota(\{a_i \in V(X) \mid i \in n\})) = M(\{\iota(a_i) \in V(X) \mid i \in n\}_{\mathcal{U}}) \\ &= \{M(\iota(a_i)) \in V(X) \mid i \in n\} \end{aligned}$$

□

Corollary 1.1.22. *The function* * *has the property*

$$ {}^*\langle a_i \rangle_i^n = \langle {}^*a_i \rangle_i^n \qquad \forall a_i \in V(X), i \in n \qquad \text{(monomorphism of tuples)} $$

Proof. This is naturally just a consequence of the fact of the monomorphism of finite sets property. Firstly, note that ${}^*\langle\rangle = {}^*\{\} = \{\}$, since $*$ takes finite sets to finite sets of equal size and there is only one set with no elements. Now by induction suppose it holds up to n-tuples. Then let $a_i \in V(X)$, $i \in n+1$ then

$$\begin{aligned} {}^*\langle a_i \rangle_i^{n+1} &= {}^*\{\{a_i \mid i \in n+1\}, \langle a_i \rangle_i^k \mid k \in n\} \\ &= \{\{{}^*a_i \mid i \in n+1\}, {}^*\langle a_i \rangle_i^k \mid k \in n\} \\ &= \{\{{}^*a_i \mid i \in n+1\}, \langle {}^*a_i \rangle_i^k \mid k \in n\} \qquad \text{(by induction hypothesis)} \\ &= \langle {}^*a_i \rangle_i^{n+1}. \end{aligned}$$

□

The next to last ingredient is;

Definition 1.1.23 (Bounded Language)**.** The bounded formulas on a set X consist of the formulas recursively defined as follows;

- The *atomic formulas on* X consist of the atomic formulas of our logic which feature only variables symbols or elements of the superstructure $V(x)$. Specifically they are of the form $x \in y$, $x = y$, $\langle x_1, \dots, x_n \rangle \in y$, and $\langle x_1, \dots, x_n \rangle = y$ where $x_1, \dots, x_n, y$ are either elements of $V(X)$ or variable symbols.

- Given any two formulas Φ, Ψ on X we can combine them in the usual manner, i.e. the following are also formulas on X: $\Phi \to \Psi$, $\Phi \wedge \Psi$, $\Phi \vee \Psi$, $\neg\Phi$, and $\Phi \leftrightarrow \Psi$.

- Given any formula Φ on X with the free variables $x_i, i \in n$ the formulas $\exists x_i(x_i \in x \wedge \Phi)$ and $\forall x_i(x_i \in x \to \Phi)$ where $i \in n$ and x is an element of $V(X)$ are formulas on X. We will in general employ the shorthands

$$\exists(x \in y)\Phi := \exists x(x \in y \wedge \Phi),$$

$$\forall(x \in y)\Phi := \forall x(x \in y \to \Phi).$$

A *sentence on* X is a formula on X which has no free variables. The sentences on X are denoted by $\mathcal{L}_X$.

And lastly we need the notion of *monomorphism*.

Definition 1.1.24 (Monomorphism)**.** A mapping $F : V(X) \to V(Y)$ from one superstructure $V(X)$ to another $V(Y)$ is called a monomorphism if it induces a "function" $\mathcal{L}_F : \mathcal{F}_X \to \mathcal{F}_Y$ from the formulas on X to the formulas on Y. Which acts as follows:

- if x is a variable symbol then $\mathcal{L}_F(x) \mapsto x$,
- if x is an element of $V(X)$ then $\mathcal{L}_F(x) \mapsto F(x)$
- for atomic formulas we have $\mathcal{L}_F(x \in y) \mapsto \mathcal{L}_F(x) \in \mathcal{L}_F(y)$ and $\mathcal{L}_F(x = y) \mapsto \mathcal{L}_F(x) = \mathcal{L}_F(y)$.
- for formulas $\Phi\Box\Psi$, where $\Box \in \{\vee, \wedge, \to, \leftrightarrow\}$, we have $\mathcal{L}_F(\Phi\Box\Psi) \mapsto \mathcal{L}_F(\Phi)\Box\mathcal{L}_F(\Psi)$,
- for formulas $\Phi = \neg\Psi$ we have $\mathcal{L}_F(\neg\Psi) \mapsto \neg\mathcal{L}_X(\Psi)$,
- and finally for formulas of the form $\Box(x \in y)\Psi$, where $\Box \in \{\exists, \forall\}$, we have $\mathcal{L}_F(\Box(x \in y)\Psi) \mapsto \Box(\mathcal{L}_F(x) \in \mathcal{L}_F(y))\mathcal{L}_F(\Psi)$.

Finally, it satisfies a/the *transfer principle*,

$$\Phi \Leftrightarrow \mathcal{L}_F(\Phi). \tag{1.4}$$

In other words, any sentence Φ on X is true if and only if $\mathcal{L}_F(\Phi)$ is true.

Unsurprisingly, we will now prove that $*$ is a monomorphism. The definition of ${}^*\Phi$ for a formula Φ is defined by the exact same scheme presented in the definition of a monomorphism. It remains only to be shown that $*$ satisfies the transfer principle;

Theorem 1.1.25 (Upward Transfer Principle)**.** *For any sentence* $\Phi \in \mathcal{L}_X$*,*

$$\Phi \Rightarrow {}^*\Phi.$$

Proof. We prove the proposition by induction on the finite construction of a sentence.

If Φ is an atomic sentence of the form: $x \in y$, $x = y$, $\langle x_1, \dots, x_n \rangle \in y$, $\langle x_1, \dots, x_n \rangle = y$, $\langle\langle x_1, \dots, x_n \rangle, x\rangle \in y$, or $\langle\langle x_1, \dots, x_n \rangle, x\rangle \in y$, then – for any choice of $x_1, \dots, x_n, x, y \in V(X)$ – the corresponding transferred sentence is true, by the monomorphism of tuple, equality, and membership properties shown in Proposition 1.1.21 and Corollary 1.1.22.

In case a sentence $\Phi = \Psi \wedge \Theta$ is true then it is true iff both Ψ and Θ are true. Now ${}^*\Psi$ and ${}^*\Phi$ are true, by the induction hypothesis. And thus, ${}^*\Psi \wedge {}^*\Theta = {}^*(\Psi \wedge \Theta)$ is true. Similarly, for the cases where we have $\vee, \rightarrow, \leftrightarrow$ instead of $\wedge$. As well as when $\Phi = \neg\Psi$.

In case a sentence $\Phi = \exists x \in y\Psi(x) \in \mathcal{L}_X$ is true then there exists an $a \in y$ such that $\Psi(a)$ is true. Thus, ${}^*\Psi(a) = ({}^*\Psi)({}^*a)$ is true in $\mathcal{L}_{{}^*X}$, by the induction hypothesis. In turn this implies that $\exists x \in {}^*y{}^*\Psi(x)$ is true in $\mathcal{L}_{{}^*X}$, since ${}^*a \in {}^*y$.

Finally, $\Phi = \forall x \in y\Psi(x)$ is true iff $\neg\exists x \in y\neg\Psi(x)$ is true and thus reduces to the previous cases. □

Theorem 1.1.26 (Downward Transfer Principle)**.** *For any sentence in $\Phi \in \mathcal{L}_X$,*

$$ {}^*\Phi \Rightarrow \Phi. $$

Proof. Let $\Phi \in \mathcal{L}_X$. Suppose $\neg\Phi$ is true, then ${}^*\neg\Phi$, by the Upward Transfer Principle (1.1.25). Moreover ${}^*\neg\Phi$ is $\neg{}^*\Phi$, by definition. □

Together the two principles form the transfer principle.

Theorem 1.1.27 (Transfer Principle)**.** *For each sentence $\Phi \in \mathcal{L}_X$, Φ is true if and only if ${}^*\Phi$ is true, i.e.*

$$ \Phi \Leftrightarrow {}^*\Phi. $$

It turns out to be useful to apply the transfer principle "partially" at times. To define precisely what we mean by this we need to define a slightly different transfer function.

Definition 1.1.28 (partial transfer function)**.** Let $\Phi \in \mathcal{L}_X$ be a formula on X with $n \in \mathbb{N}$ free variables then the partial transfer ${}^\dagger\Phi$ of Φ is the formula in ${}^*\Phi$ except the each free variable x_i is substituted with the expression ${}^*x_i$.

We can now state the partial transfer principle;

Theorem 1.1.29 (partial transfer principle)**.** *Let $\Psi \in \mathcal{L}_X$ be a sentence on X with a subformula $\Phi \in \mathcal{L}_X$ and let Θ be the sentence Ψ with Φ substituted with ${}^\dagger\Phi$. Then,*

$$ \Psi \Leftrightarrow \Theta. $$

Proof. Let Ψ, Φ, and Θ be as in the proposition statement. Consider first the formula $\Phi \in \mathcal{L}_X$ suppose it has $n \in \mathbb{N}$ free variables x_i, $i \in n$. We denote this by $\Phi(x_i; i \in n)$. Now, for any choice of constants $b_i \in V(X)$ we have that

$$ ({}^\dagger\Phi)(b_i; i \in n) \Leftrightarrow ({}^*\Phi)({}^*b_i; i \in n) \Leftrightarrow {}^*(\Phi(b_i; i \in n)) \Leftrightarrow \Phi(b_i; i \in n), $$

by definition and the Transfer Principle (1.1.27). Since, $\Psi \in \mathcal{L}_X$ all the bounds of the free variables x_i in Φ constrain the values of x_i to elements of $V(X)$. The theorem follows. □

We note that the partial transfer principle is not found in this explicit form anywhere in the literature. At least to our knowledge. We include it because it turns deriving many nonstandard characterisations of standard notions into a completely mechanical excercise in symbolic manipulation.

1.2 The Transfer Principle

Let us consider some of the direct consequences and nomenclature that relates to the transfer principle. One immediate consequence is;

Proposition 1.2.1 (∗-commutation)**.** *For any $M \in V(X)$ and $\Phi \in \mathcal{L}_X$ the following commutation property holds*

$$^*\{x \in M \mid \Phi(x)\} = \{x \in {}^*M \mid {}^*\Phi(x)\}.$$

Proof. Let M and Φ be as in the statement of the proposition. Now the equivalence

$$\forall x \in M(x \in \{z \in M \mid \Phi(z)\} \leftrightarrow \Phi(x)) \Leftrightarrow \forall x \in {}^*M(x \in {}^*\{z \in M \mid \Phi(z)\} \leftrightarrow {}^*\Phi(x)),$$

holds, by the Transfer Principle (1.1.27). □

Any definition that can be expressed in first order logic on our set X has a related "∗-notion" (read "star-notion"). On the simpler side we call $^*\mathbb{N}$ the ∗-natural numbers, $^*\mathbb{R}$ the ∗-real numbers, $^*\mathbb{Q}$ the ∗-rational numbers. For a slightly more elaborate example, recall that the convergent $\mathbb{R}$ sequences consist of the elements of the set

$$A := \{x \in \mathbb{R}^{\mathbb{N}} \mid \exists r \in \mathbb{R}\forall \varepsilon > 0\exists n \in \mathbb{N}\forall m \in \mathbb{N} + n(|r - x_m| < \varepsilon)\}.$$

(Note that the equality follows from ∗-commutation (1.2.1).) Consequently, the ∗-convergent $^*\mathbb{R}$ sequences consist of the elements of

$$^*A = \{x \in {}^*(\mathbb{R}^{\mathbb{N}}) \mid \exists r \in {}^*\mathbb{R}\exists n \in {}^*\mathbb{N}\forall m \in {}^*\mathbb{N} + n(|r - x_m| < \varepsilon)\}.$$

For a slightly less trivial notion consider the following, which we shall need repeatedly.

Definition 1.2.2 (∗-finite)**.** The ∗-finite sets are the elements of

$$\bigcup_n^{\mathbb{N}}\{A \in {}^*V_n(X) \mid {}^*|A| \in {}^*\mathbb{N}\} = \bigcup_n^{\mathbb{N}} {}^*\{A \in V_n(X) \mid |A| \in \mathbb{N}\}. \tag{1.5}$$

(Note, again, that the equality above follows from ∗-commutation (1.2.1).) We see that the ∗-finite sets consist of the union of the ∗-value of the finite sets in each $V_n(X)$. The full set of finite sets in $V(X)$ is not an element of $V(X)$ and so we cannot define ∗-finiteness in terms the value * assigns to the set of finite sets in $V(X)$. Note that, by the transfer principle, every first order property of a set A in $V(X)$ also holds for $^*A$. For instance, if A is ∗-finite then a ∗-function $f : A \to A$ is a ∗-injection if and only if it is a ∗-surjection. Note that we did not say that any function $f : A \to A$ is an injection if and only if it is a surjection! Since, the properties that hold for ∗-notions are the ∗-transferred ones. We will return to this in section 1.3 Standard, Nonstandard, Internal, and External objects.

Prefixing a regular notion by "∗-" is a general naming convention in nonstandard analysis. Some authors use the "hyper" prefix instead of the "∗-" prefix, i.e. they call $^*\mathbb{R}$ the hyperreal numbers, $^*\mathbb{N}$ the hypernatural numbers, etc. In this text we will not, in general, use the "hyper" style terminology in this way. We choose to prefix the transferred notion by "∗-".

Next we look at some more commutation properties of the transfer function.

Proposition 1.2.3. *Let $C, B, A_i \in V(X)$ for $i \in n \in \mathbb{N}$ then $^*\left(\prod_i^n A_i\right) = \prod_i^n {}^*A_i$, $^*\left(\bigcup_i^n A_i\right) = \bigcup_i^n {}^*A_i$, $^*\left(\bigcap_i^n A_i\right) = \bigcap_i^n {}^*A_i$, and $^*(C \setminus B) = {}^*C \setminus {}^*B$. In other words, ∗ commutes with finite (cartesian product), finite union, finite intersection, and set difference.*

Proof. The proof is again by transfer. We prove the case for the cartesian product. Let A_i, n be as in the statement of the proposition. Since

$$\forall z \in \prod_i^n A_i \forall i \in n \exists z_i \in A_i((z_i)_i^n = z) \Leftrightarrow \forall z \in {}^*\left(\prod_i^n A_i\right) \forall i \in n \exists z_i \in {}^*A_i((z_i)_i^n = z)$$

it follows that ${}^*\prod_i^n A_i \subseteq \prod_i^n {}^*A_i$. And since

$$\forall i \in n \forall z_i \in A_i \exists z \in \prod_i^n A_i((z_i)_i^n = z) \Leftrightarrow \forall i \in n \forall z_i \in {}^*A_i \exists z \in {}^*\prod_i^n A_i((z_i)_i^n = z)$$

it follows that $\prod_i^n {}^*A_i \subseteq {}^*\prod_i^n A_i$. And so $\prod_i^n {}^*A_i = {}^*\prod_i^n A_i$.

The proofs for the remaining properties is similar. □

It should be obvious why the finiteness of the operations is essential in the above theorem.

One notable set operation with which * does not commute in general is ${}^*X^{{}^*Y} \neq {}^*\left(X^Y\right)$ for sets $X, Y \in V(X)$, when Y is infinite. We will prove the this inequality in section 1.3 Standard, Nonstandard, Internal, and External objects.

As a final note on the transfer principle we give an example of how the partial transfer principle is usually employed. It is frequently useful when deriving the nonstandard characterisation of some notion.

We take as our example a derivation of the nonstandard characterisation of continuity in a metric space. In the following $x \simeq 0$ is defined to mean that $x < \varepsilon$ for all $\varepsilon \in \mathbb{R}_{>0}$. Furthermore we write $x \simeq y$ if $d(x, y) \simeq 0$,

$${}^\sigma A = {}^*[A] = \{{}^*a \mid a \in A\},$$

and we assume that there exists $\delta \in {}^*\mathbb{R}_{>0}$ such that $\delta \simeq 0$ (i.e. that our extension has infinitesimals). (We will return to all these notions in greater detail later.)

Proposition 1.2.4. *A function $f : X \to Y$ is continuous if and only if*

$$\forall x \in {}^\sigma X \forall y \in {}^*X({}^*d(x, y) \simeq 0 \to {}^*d({}^*f(x), {}^*f(y)) \simeq 0).$$

Proof. Starting from the usual definiton of continuity, consider the following sequence of implications. (Recall that † stands for the partial transfer function.)

$$\begin{aligned}
&\forall x \in X \forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in \mathbb{R}_{>0} \forall y \in X(d(x, y) < \delta \ \to\ d(f(x), f(y)) < \varepsilon)\\
&\Leftrightarrow \forall x \in X \forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in \mathbb{R}_{>0}{}^\dagger(\forall y \in {}^*X({}^*d(x, y) < \delta \to {}^*d({}^*f(x), {}^*f(y)) < \varepsilon))\\
&\Leftrightarrow \forall x \in X \forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in \mathbb{R}_{>0} \forall y \in {}^*X({}^*d({}^*x, y) < {}^*\delta \to {}^*d({}^*f({}^*x), {}^*f(y)) < {}^*\varepsilon)\\
&\Leftrightarrow \forall x \in {}^\sigma X \forall \varepsilon \in {}^\sigma\mathbb{R}_{>0} \exists \delta \in {}^\sigma\mathbb{R}_{>0} \forall y \in {}^*X({}^*d(x, y) < \delta \to {}^*d({}^*f(x), {}^*f(y)) < \varepsilon)\\
&\Rightarrow \forall x \in {}^\sigma X \forall \varepsilon \in {}^\sigma\mathbb{R}_{>0} \forall y \in {}^*X({}^*d(x, y) \simeq 0 \to {}^*d({}^*f(x), {}^*f(y)) < \varepsilon)\\
&\Leftrightarrow \forall x \in {}^\sigma X \forall y \in {}^*X({}^*d(x, y) \simeq 0 \to {}^*d({}^*f(x), {}^*f(y)) \simeq 0)
\end{aligned}$$

Conversely,

$$\begin{aligned}
&\forall x \in {}^\sigma X \forall y \in {}^*X({}^*d(x, y) \simeq 0 \to {}^*d({}^*f(x), {}^*f(y)) \simeq 0)\\
&\Rightarrow \forall x \in {}^\sigma X \exists \delta \in {}^*\mathbb{R}_{>0} \forall y \in {}^*X({}^*d(x, y) < \delta \to {}^*d({}^*f(x), {}^*f(y)) \simeq 0)\\
&\Rightarrow \forall x \in {}^\sigma X \forall \varepsilon \in {}^\sigma\mathbb{R}_{>0} \exists \delta \in {}^*\mathbb{R}_{>0} \forall y \in {}^*X({}^*d(x, y) < \delta \to {}^*d({}^*f(x), {}^*f(y)) < \varepsilon)\\
&\Leftrightarrow \forall x \in X \forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in {}^*\mathbb{R}_{>0} \forall y \in {}^*X({}^*d({}^*x, y) < \delta \to {}^*d({}^*f(x), {}^*f(y)) < {}^*\varepsilon)\\
&\Leftrightarrow \forall x \in X \forall \varepsilon \in \mathbb{R}_{>0}{}^\dagger(\exists \delta \in \mathbb{R}_{>0} \forall y \in X(d(x, y) < \delta \to d(f(x), f(y)) < \varepsilon))\\
&\Leftrightarrow \forall x \in X \forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in \mathbb{R}_{>0} \forall y \in X(d(x, y) < \delta \to d(f(x), f(y)) < \varepsilon).
\end{aligned}$$

□

In general we will suppress the $^\sigma$ superscripts. In particular, from this point on, whenever we refer to a set $A \in V(X)$ as though it were in $V({}^*X)$ we are in fact referring to ${}^\sigma A$, not ${}^*A$. We adopt this convention solely to reduce visual clutter. Furthermore we will also, for the most part, omit the * superscript except on the bounds of quantifiers. (As we already did in the above proof for the $<$ relation, where we should technically have written '${}^* <$'.) Moreover, the step where we remove the *'s from the variables by applying $^\sigma$ to the set bounding it may be skipped in future proofs.

1.3 Standard, Nonstandard, Internal, and External objects

When making arguments there will be a need to distinguish between different classes of object in $V({}^*X)$.

Recall the notation

$$ {}^\sigma A := {}^*[A] = \{ {}^*a \mid a \in A \} $$

for the $*$-image of elements $A \in V(X) \setminus X$. (We will use the square bracket notation for the image later on as well. In general, if f is a function and A a set in the domain of f we denote the f-image of A by $f[A]$.) We are not yet in a position to prove that ${}^\sigma A$ and ${}^*A$ are not the same set in general. The proof is provided in section 1.4 Saturation Principles. As already mentioned, when the context is clear we will suppress the ${}^\sigma$. In particular, when we write $\mathbb{R}$ or $\mathbb{N}$, the context will often decide which one we referred to.

Definition 1.3.1 (Standard & Nonstandard)**.** An object is *standard* if it is an element of

$$ {}^\sigma V(X) := \bigcup_n^{\mathbb{N}} {}^\sigma V_n(X). $$

An object is *nonstandard* if it is not standard.

A formula on $V({}^*X)$, i.e. in $\mathcal{L}_{{}^*X}$, is standard if the constants in it are standard, otherwise it is nonstandard.

Crucially, only the standard sentences can be transferred back to $V(X)$. The following is arguably the most important dichotomy in the nonstandard framework;

Definition 1.3.2 (Internal & External)**.** We define the set of *internal* elements of $V(X)$ as the set

$$ {}^*V(X) = \bigcup_n^{\mathbb{N}} {}^*V_n(X). $$

The set of *external* objects is $V({}^*X) \setminus {}^*V(X)$.

A formula on $V({}^*X)$ is internal if the constants in it are internal, otherwise it is external.

The reason internal objects, in particular, are important is that they are, loosely speaking, the objects which satisfy the transferred versions of the properties we are used to. For instance, the internal subsets of ${}^*\mathbb{R}$ which have an upper bound in ${}^*\mathbb{R}$ have a unique supremum in ${}^*\mathbb{R}$. Note that every standard object is also internal, while some nonstandard objects are internal and some are external.

We can also restate the definition of $*$-finite sets in terms of the internal objects ${}^*V(X)$;

Definition 1.3.3 ($*$-finite)**.** The $*$-finite sets consist of the elements of

$$ \{ A \in {}^*V(X) \mid {}^*|A| \in {}^*\mathbb{N} \}. $$

Moreover we define the $*$-functions as follows;

Definition 1.3.4 ($*$-function)**.** The $*$-functions consist of the elements of

$$ \left\{ f \in {}^*V(X) \middle| \exists A, B \in {}^*V(X) (f \in A^B) \right\}. $$

Note that by definition the $*$-functions are functions, since they are elements of A^B for some internal sets A, B. But, not every element of A^B will in general be a $*$-function, since not every element of A^B need be internal.

We can therefore observe that in general when we refer to an object as being $*$-finite, a $*$-sequence, etc. we are also saying that the object in question is internal. Or, at any rate, that is how it should be. In practice, it is so natural to think of a function $f : {}^*\mathbb{N} \to X$ as a

$*$-sequence that we will not always rely on this convention. Therefore, while it will at times be assumed that an object is internal if it is referred to as a $*$-'something' (e.g. $*$-finite), we will not shy away from also stating explicitly that it is internal for clarity. For instance, you should assume that an object is internal if it is referred to as being $*$-finite. On the other hand, we will, in general, state explicitly that a $*$-sequence is internal.

As we mentioned previously, the internal objects are the ones that the transferred sentences are about; And which therefore behave "normally". Let us elaborate on the example of internal subsets of ${}^*\mathbb{R}$ with an upper bound having a supremum. Formally we have the following from standard analysis:

$$\forall A \in \mathcal{P}(\mathbb{R})(\exists r \in \mathbb{R} \forall s \in A(s < r) \to \exists! r \in \mathbb{R}(\forall s \in \mathbb{R}(s < r) \wedge \forall t \in \mathbb{R}_{<r} \exists s \in A(t < s)).$$

It follows that the internal subsets of ${}^*\mathbb{R}$ with an upper bound have a unique supremum (in ${}^*\mathbb{R}$). Since,

$$\forall A \in {}^*\mathcal{P}(\mathbb{R})(\exists r \in {}^*\mathbb{R} \forall s \in A(s < r) \to \exists! r \in {}^*\mathbb{R}(\forall s \in {}^*\mathbb{R}(s < r) \wedge \forall t \in {}^*\mathbb{R}_{<r} \exists s \in A(t < s)),$$

by transfer. (We show in Proposition 1.3.5 the elements of ${}^*\mathcal{P}(\mathbb{R})$ are precisely the internal subsets of ${}^*\mathbb{R}$.) On the other hand, external subsets of ${}^*\mathbb{R}$ need not have a unique supremum in ${}^*\mathbb{R}$ despite having an upper bound in ${}^*\mathbb{R}$. An example of such an external set is ${}^\sigma\mathbb{R}$. Any infinite element of ${}^*\mathbb{R}$, i.e. larger than any standard element of ${}^*\mathbb{R}$, is an upper bound of ${}^\sigma\mathbb{R}$. Suppose R is the supremum of ${}^\sigma\mathbb{R}$ in ${}^*\mathbb{R}$. But now, $R-1$ is also an upper bound of ${}^\sigma\mathbb{R}$ since it is also infinite. Which contradicts our assumption that R was the least upper bound of ${}^\sigma\mathbb{R}$ in ${}^*\mathbb{R}$. (To see that $R-1$ must also be infinite note that if it is not, thenn it is smaller than some standard $a \in {}^\sigma\mathbb{R}$. That is $R-1 < a$, but now $R < a+1$ and $(a+1) \in {}^\sigma\mathbb{R}$, which contradicts the assumption that R is larger than every element of ${}^\sigma\mathbb{R}$. The existence of infinite $*$-real numbers requires a saturation argument and is therefore postoned.) Indeed the lack of a supremum in ${}^*\mathbb{R}$ for ${}^\sigma\mathbb{R}$ despite being bounded in ${}^*\mathbb{R}$ proves that ${}^\sigma\mathbb{R}$ is external.

This is an instance of a general pattern for showing that a set is external. Namely, that it lacks a property that all internal sets have by transfer. We will return to these considerations later.

Proposition 1.3.5. *The internal subsets of an internal set A are precisely the elements of ${}^*\mathcal{P}(A)$.*

Proof. Consider,

$$\begin{aligned}
&\forall A, B \in V(X)(\forall s \in B(s \in A) \leftrightarrow B \in \mathcal{P}(A) \leftrightarrow B \subset A) \\
&\Leftrightarrow \forall n \in \mathbb{N} \forall A, B \in V_n(X)(\forall s \in B(s \in A) \leftrightarrow B \in \mathcal{P}(A) \leftrightarrow B \subset A),
\end{aligned}$$

by transfer,

$$\begin{aligned}
&\Leftrightarrow \forall n \in \mathbb{N} \forall A, B \in {}^*V_n(X)(\forall s \in B(s \in A) \leftrightarrow B \in {}^*\mathcal{P}(A) \leftrightarrow B^* \subset A) \\
&\Leftrightarrow \forall A, B \in {}^*V(X)(\forall s \in B(s \in A) \leftrightarrow B \in {}^*\mathcal{P}(A) \leftrightarrow B^* \subset A).
\end{aligned}$$

□

The above proof suggests that $B^* \subset A$ only makes sense if both A and B are internal and B is a subset of A. When we write $B \subset A$ for a set $A \in V(X)$ we mean that every element of B is in A. This, naturally, says nothing about whether A or B are internal. However, $B^* \subset A$ is equivalent to saying that $B \in {}^*\mathcal{P}(A)$ and the operation ${}^*\mathcal{P}$ is only defined for internal sets, i.e. sets $A \in {}^*V(X)$. Thus, there is an important difference between $\mathcal{P}$ and ${}^*\mathcal{P}$; Not only are the elements of the latter internal, but ${}^*\mathcal{P}$ can only be applied on internal sets. On the other hand, it would be natural to denote the internal subsets of an external set A by ${}^*\mathcal{P}(A)$. Regardless, to avoid confusion we will not use $^* \subset$ nor $^*\mathcal{P}$ to specify that a set is internal. Instead, as we noted earlier, we will always state explicitly when something is assumed to be internal.

In the following propositions we prove essentially all the remaining fundamental facts about internal objects.

Proposition 1.3.6. *Elements of internal objects are internal.*

Proof. For each $n \in \mathbb{N}_{\geq 1}$,

$$\forall x \in (V_n(X) \setminus V_{n-1}(X)) \forall y \in x (y \in V_{n-1}(X)),$$

is true. Thus, by transfer,

$$\forall x \in {}^*\mathcal{P}(V_{n-1}(X)) \forall y \in x (y \in {}^*V_{n-1}(X)).$$

□

Theorem 1.3.7 (Keisler's internal definition principle)**.** *Let A be an internal set and let $\Phi(x)$ be an internal formula with a free variable x. Then, the set*

$$\{a \in A \mid \Phi(a)\}$$

is internal.

Proof. Let $\Psi(x)$ be an internal formula on $V({}^*X)$ with $n \in \mathbb{N}$ internal constants c_i, for each $i \in n$ and free variable x. We will write this as $\Psi(x) = \Phi(c_i, i \in n; x)$. Moreover, let A be an internal set. Then, there is some ${}^*V_m(X)$ such that $c_i \in V_m(X)$ for all $i \in n$ and $A \in {}^*V_m(X)$. Then,

$$\forall x_0, \dots, x_{n-1}, x \in V_m(X) \exists y \in V_{m+1}(X) \forall z \in V_m(X)(z \in y \leftrightarrow (z \in x \wedge \Phi(x_i, i \in n; x))) \quad (1.6)$$

is true from regular set theory and it is a formula on $V(x)$. Transferring the sentence yields the theorem. □

Proposition 1.3.8. *Internal sets are closed under the regular set operations. In particular if A, B are internal then $A \cap B$, $A \setminus B$, $A \cup B$, and $A \times B$ are internal.*

Proof. The proof is by a direct application of the transfer principle. □

The following sets will be used repeatedly;

Definition 1.3.9 ($\mathbb{N}_\infty$,$\mathbb{R}_\infty$,$\mathbb{R}_{\simeq 0}$)**.**

$$\begin{aligned} \mathbb{N}_\infty &= \{n \in {}^*\mathbb{N} \mid \forall m \in \mathbb{N}(m < n)\} \\ \mathbb{R}_\infty &= \{r \in {}^*\mathbb{R} \mid \forall s \in \mathbb{R}(|s| < |r|)\} \\ \mathbb{R}_{\simeq 0} &= \{r \in {}^*\mathbb{R} \mid \forall s \in \mathbb{R}(|r| < |s|)\} \end{aligned}$$

We call them the infinite $*$-naturals, infinite $*$-real numbers, and infinitesimal $*$-real numbers respectively.

This sort of notation is extended to the integers, the extended reals, etc. in the natural way. To prove that $\mathbb{N}_\infty$ and $\mathbb{R}_\infty$ are non-empty and that $\mathbb{R}_{\simeq 0}$ contains more than just 0 requires a saturation argument, which is provided in section 1.4 Saturation Principles. (Technically we would need to assume that * is an enlargement (1.4.2).) In the following proof the non-emptyness is simply assumed.

Proposition 1.3.10. *The sets $\mathbb{N}, \mathbb{R}, \mathbb{N}_\infty$,$\mathbb{R}_\infty, \mathbb{R}_{\simeq 0}$ are external.*

Note that here $\mathbb{N}$ and $\mathbb{R}$ refer to ${}^\sigma\mathbb{N}$ and ${}^\sigma\mathbb{R}$ respectively. As indicated by the context, since we state that they are external, that is they are elements of $V({}^*X) \setminus {}^*V({}^*X)$, and thus are elements of $V({}^*X)$ which is precisely when the ${}^\sigma$ is implicit. (Recall that X stands for the set of atomic entities we want to reason about using nonstandard arguments.)

Proof. By transfer any internal subset of ${}^*\mathbb{R}$ with an upper bound (lower buound) must have a unique supremum (infimum) in ${}^*\mathbb{R}$. But, the elements of $\mathbb{N}_\infty$ are upper bounds of $\mathbb{N}$, yet no element $\kappa \in \mathbb{N}_\infty$ can be the supremum of $\mathbb{N}$ since $\kappa - 1$ is also in $\mathbb{N}_\infty$ and thus an upper bound of $\mathbb{N}$. Similarly for $\mathbb{R}$ and $\mathbb{R}_{\simeq 0}$. Next, note that the elements of $\mathbb{N}$ are lower bounds for $\mathbb{N}_\infty$. Yet no element $n \in \mathbb{N}$ can be the infimum of $\mathbb{N}_\infty$, since $n+1$ is also an element $\mathbb{N}$ and thus a lower bound of $\mathbb{N}_\infty$. A similar argument applies to $\mathbb{R}_\infty$. □

Finally, we have the following transfer of the induction principle;

Proposition 1.3.11 ($*$-induction)**.** *For any internal formula $\Phi(x)$, with free variable x, if $\Phi(0)$ is true and $\Phi(n) \to \Phi(n+1)$ for each $n \in {}^*\mathbb{N}$ then $\forall n \in {}^*\mathbb{N}\Phi(n)$ holds.*

Proof. Let $\Phi(c_i, n \mid i \in m)$ be an internal formula with $m \in \mathbb{N}$ internal constants c_i. Then there exists some $l \in \mathbb{N}$ such that $c_i \in {}^*V_l(X)$ for each $i \in m$. Now the formula

$$\Psi(n) := \forall x_0, \ldots, x_{m-1} \in {}^*V_l(X)\Phi(x_i, n \mid i \in m)$$

is standard. Let Θ be the unique formula such that ${}^*\Theta = \Psi$. Observe that Θ is unique since the transfer function is an injection and exists since Ψ is standard. By the principle of induction it follows that

$$(\Theta(0) \wedge \forall n \in \mathbb{N}(\Theta(n) \to \Theta(n+1))) \to \forall n \in \mathbb{N}\Theta(n)$$

and, by transfer,

$$({}^*\Theta(0) \wedge \forall n \in {}^*\mathbb{N}(\Theta(n) \to \Theta(n+1)) \to \forall n \in {}^*\mathbb{N}\Theta(n).$$

We introduce the following shortand: for any formula $\Gamma(x)$ with free variable x, let

$${}^*P(\Gamma) := (\Gamma(0) \wedge \forall n \in {}^*\mathbb{N}(\Gamma(n) \to \Gamma(n+1))) \to \forall n \in {}^*\mathbb{N}\Gamma(n).$$

The conclusion now follows from the following sequence of implications;

$$\begin{aligned} P({}^*\Theta) &\Rightarrow P(\Psi) \\ &\Rightarrow \forall x_0, \ldots, x_m \in {}^*V_l(X)P(\Phi(x_i \mid i \in m)) \\ &\Rightarrow P(\Phi(c_i \mid i \in m)) \\ &\Rightarrow (\Phi(0) \wedge \forall n \in {}^*\mathbb{N}(\Phi(n) \to \Phi(n+1))) \to \forall n \in {}^*\mathbb{N}\Phi(n). \end{aligned}$$

□

1.4 Saturation Principles

The saturation of $V(^*X)$, roughly speaking, refers to how many new objects are in the $*$-transfer of elements of $V(X)$. Technically, saturation is a property of the transfer function * since we are talking about how it relates $V(X)$ and $V(^*X)$.

The most important and powerful form of saturation property is κ-saturation (1.4.8). The other saturation property, enlargement (1.4.2), that we cover is strictly weaker. Any necessary degree of saturation can always be achieved. However, the proofs for this are beyond the scope of this work. In general it is a matter of choosing a sufficiently large index set and a sufficiently carefully constructed ultrafilter.

We will in general not state the necessary saturation properties that our model needs in proposition statements. This is mainly because the saturation required is not of interest from the point of view of doing analysis.

The notion of $*$-finite plays a role in saturation. Recall the following definition.

Definition 1.4.1 ($*$-finite)**.** If A is an internal subset of $V(^*X)$, then the $*$-*finite* subsets of A are the elements of $^*\mathcal{P}_F(A)$.

The first and weakest form of saturation we will discuss is;

Definition 1.4.2 (enlargement)**.** The mapping $* : V(X) \to V(^*X)$ is an enlargement if for every set $A \in V(X)$ there exists a set $B \in {}^*\mathcal{P}_F(A)$ such that $^\sigma A \subset B$.

In a sense this is just to say that the sets, in $V(X)$ are all small enough in our extension that they are subsets of some $*$-finite set. However, they are not necessarily $*$-finite themselves.

In Robinson's original formulation he defined enlargement in terms of concurrent relations. The relation between these two notions will frequently be useful.

Definition 1.4.3 (Concurrent Relation)**.** A relation $R \subset A \times B$ is called *concurrent* (aka. *finitely satisfiable*) if for any finite subset $S \subset B$ there is an element $x \in A$ such that xRs for all $s \in S$. If xRs for all $s \in S$ we say that x R-satisfies S. In these terms the definition can be rephrased as: the relation $R \subset A \times B$ is *concurrent* if every finite subset of B can be R-satisfied by an element of A.

The two notions are related by the following proposition.

Proposition 1.4.4 (Enlargement & Concurrency)**.** *The correspondence $* : V(X) \to V(^*X)$ is an enlargement if and only if for every concurrent relation $R \subset A \times B$ in $V(X)$ there is an element $x \in {}^*A$ that $^*R$-satisfies $^\sigma B$.*

Proof. Let $V(^*X)$ be an enlargement and $R \subset A \times B \in V(X)$ be a concurrent relation. Then, for each $*$-finite subset S of $^*B$ there is an element $x \in {}^*A$ such that $x{}^*Ry$ for all $y \in S$. Thus, this also applies to any $C \in {}^*\mathcal{P}_F(B)$ such that $C \supset {}^\sigma B$. Since at least one such set exists, by enlargement, it follows that there is at least one element of $^*A$ which $^*R$-satisfies all elements of $^\sigma B$.

Conversely, let $A \in V(X)$. Now consider the relation

$$R := \{(x, y) \in \mathcal{P}_F(A) \times A \mid y \in x\}.$$

This relation is obviously concurrent. Furthermore $\mathcal{P}_F(A)$ is an element of $V(X)$, since $A \in V(X)$. It follows that there is some element $B \in {}^*\mathcal{P}_F(A)$ such that $(x, B) \in {}^*R$ for all $x \in {}^\sigma A$. That is $B \in^* \mathcal{P}_F(A)$ and $x \in B$ for all $x \in {}^\sigma A$. The proposition follows. ☐

Proposition 1.4.5 (existence of nonstandard objects)**.** *In enlargements the $*$-transfer of any infinite set contains nonstandard elements.*

Proof. Let $A \in V(X)$ be an infinite set. Then the relation

$$R := \{(x, y) \in A \times A \mid x \neq y\},$$

is finitely satisfiable. Therefore there is an element $b \in {}^*\mathcal{P}_F(A)$ such that $b \neq a$ for all $a \in {}^\sigma A$, by Proposition 1.4.4. □

An immediate corollary is that ${}^*A$ and ${}^\sigma A$ are not in general the same set. In particular;

Corollary 1.4.6. *If $A \in V(X)$ is an infinite set and * is an enlargement, then ${}^\sigma A \neq {}^*A$.*

Proof. There exists an element x in ${}^*A$ such that $x \neq {}^*a$ for all $a \in A$, by the existence of nonstandard objects (1.4.5). In other words, $x \neq a$ for all $a \in {}^\sigma A$. □

Another consequence, which we have alluded to earlier, is the following;

Proposition 1.4.7. *If * is an enlargement, then the sets $\mathbb{N}_\infty$, $\mathbb{R}_\infty$ are non-empty and $\mathbb{R}_{\simeq 0}$ is a strict superset of $\{0\}$.*

Proof. We begin by proving that there are non-zero $*$-real infinitesimals. By enlargement, there exists a $*$-finite set $A \subset {}^*(0,1]$ such that ${}^\sigma(0,1] \subset A$. Now, since the $*$-real numbers are closed under $*$-finite products, it follows, by the usual rules of arithmetic, that for each $b \in (0,1]$ one has

$$0 < \prod_a^A a < b.$$

Consequently, $\prod_a^A a \in \mathbb{R}_{\simeq 0}$ and $\prod_a^A a \neq 0$.

Next we show that $\mathbb{N}_\infty$ is non-empty. By enlargement, there exists a $*$-finite set $A \subset {}^*\mathbb{N}$ such that ${}^\sigma\mathbb{N} \subset A$. Now, since $*$-natural numbers are closed under $*$-finite addition, it follows, by the usual rules of arithmetic, that for each $m \in {}^\sigma\mathbb{N}$

$$m < \sum_n^A n.$$

Consequently, $\sum_n^A n \in \mathbb{N}_\infty$.

Finally, since $\mathbb{N}_\infty \subset \mathbb{R}_\infty$ it follows that $\mathbb{R}_\infty$ is non-empty. □

Above we find the first instance of an 'infinite summation', in the sense of: a sum of infinitely many summands. (In contrast to an 'infinite sum' which an infinitely large sum. In the proof we have both an infinite sum and an infinite summation.) Despite the infinites involved we know the internal arithmetic of ${}^*\mathbb{R}$ and ${}^*\mathbb{N}$ follows all the "usual rules" as a, fairly direct, consequence of transfer.

Finally, we present the most powerful saturation principle. In this regard it is the only one one needs to remember. Since one is, in general, not concerned with keeping track of saturation when applying nonstandard arguments to, e.g. analysis. However, in practice enlargement is often simpler to invoke when sufficient.

Definition 1.4.8 (κ-saturation). The correspondence $* : V(X) \to V({}^*X)$ is called κ-saturated if for each internal relation $R \subset A \times B \in {}^*V(X)$ concurrent on some (possibly external) set $C \subset B$ such that $\text{card}(C) < \kappa$ there is an element of A which R-satisfies C.

The following two consequences of κ-saturation turn out to cover most of the cases where κ-saturation itself is needed but the relational formulation is not the most natural. It turns out that they are equivalent to κ-saturation. In particular, the κ-intersection principle, is sometimes given instead of κ-saturation as the primary saturation principle. The case $\kappa = \aleph_1$ is perhaps its most commonly encountered form.

We will need the following definition to state the κ-intersection principle:

Definition 1.4.9 (finite intersection property)**.** A collection of sets C is said to have the *finite intersection property* if the intersection of any finite subset D of C has non-empty intersection, i.e. $\bigcap D \neq \emptyset$.

Proposition 1.4.10 (κ-intersection principle)**.** *If the correspondence * is κ-saturated, C is a collection of internal subsets of some internal set B, card$(C) < \kappa$, and C satisifes the finite intersection property, then C has non-empty intersection.*

Proof. Let C be as in the proposition and the correspondence * be κ-saturated. Internal sets are closed under intersection, by Keisler's internal definition principle. Consider the relation

$$R := \{(x, y) \in \left(\bigcap_{c}^{C} c\right) \times C \mid x \in y\}.$$

It is clearly internal and concurrent on C. It follows that there is a non-empty internal subset of B which is in the intersection of the elements of C, by κ-saturation (1.4.8). $\square$

Proposition 1.4.11 (κ-extension principle)**.** *If the correspondence * is κ-saturated, then for each function $f : A \subset B \to C$, where A is a (possibly external) subset of an internal subset B, C is internal, and card$(A) < \kappa$, there exist an internal extension $\hat{f} : B \to C$.*

Proof. Let $f : A \subset B \to C$ be as in the proposition and the model be κ-saturated. Then the relation

$$R := \{(g, (b, c)) \in (B \to C) \times (B \times C) \mid g(b) = c\},$$

is internal and concurrent on $\Gamma = \{(a, c) \in A \times f[A] \mid f(a) = c\}$. The set Γ is the graph of f and thus has the same cardinality as A. The cardinality of Γ is therefore lower than κ. It follows that there is an internal function $\hat{f} \in B \to C$ such that $\hat{f}(a) = c = f(a)$ for all $a \in A$, by κ-saturation (1.4.8). $\square$

In practice, when we use the κ-extension principle on a function we do not in general bother introducing a new name (or symbol) for the extended function. Instead once the extension principle is invoked we refer to the extension by the name of the original. (We can always restrict the domain to recover the original function if necessary.) For example, we might invoke the κ-extension principle to extend a sequence $f : \mathbb{N} \to A$ to an internal $*$-sequence $f : {}^*\mathbb{N} \to A$.

1.5 Permanence Principles

While the propositions of this section are staples of nonstandard analysis the names vary from author to author. We have tried to assign them names that seem consistent and memorable.

In general, what we call a permanence principle is any proposition that proves that if we know an internal formula $\Phi(x)$ to be true for each x in some external set A then it must hold in some internal superset set as well.

Which is essentially, precisely, what our simplest permanence principle states;

Proposition 1.5.1 (Spillover). *Suppose $\{x \in B \mid \Phi(x)\} \supseteq A$ where B is an internal set, $\Phi(x)$ is an internal formula, and A is an external set. Then $\{x \in B \mid \Phi(x)\}$ is strictly larger than A, i.e. it contains an element of B which is not in A.*

Proof. Suppose that $\{x \in B \mid \Phi(x)\} = A$, then it is an external set. However, the set $\{x \in B \mid \Phi(x)\}$ is internal, by Keisler's internal definition principle (1.3.7). It must therefore be a strict superset of A, since external sets are not internal, by definition. □

The following is called the Permanence principle by some authors and is the one we will invoke the most frequently;

Theorem 1.5.2 (Cauchy's principle). *Let * be an enlargement and Φ be an internal formula. Then the following hold;*

1. *If $\Phi(x)$ is true for all $x \in \mathbb{N}$, then it is true up to some $k \in \mathbb{N}_\infty$.*

2. *If $\Phi(x)$ is true for all $x \in \mathbb{N}_\infty$, then it is true down to some $k \in \mathbb{N}$.*

3. *If $\Phi(x)$ is true for all positive infinitesimals $x \in {}^*\mathbb{R}$, then it is true up to some $x \in {}^\sigma\mathbb{R}$.*

Recall the definitions of the sets.

$$\mathbb{N}_\infty := \{n \in {}^*\mathbb{N} \mid \forall m \in \mathbb{N}(m < n)\}$$
$$\mathbb{R}_\infty := \{r \in {}^*\mathbb{R} \mid \forall s \in \mathbb{R}(s < r)\}$$
$$\mathbb{R}_{\simeq 0} := \{r \in {}^*\mathbb{R} \mid \forall s \in \mathbb{R}_{>0}(|r| < s)\}.$$

Proof. Firstly, by enlargement, $\mathbb{N}_\infty$, $\mathbb{R}_\infty$, and $\mathbb{R}_{\simeq 0}$ are nonempty.

Now observe that by transfer any internal subset of $^*\mathbb{R}$ which has an upper or lower bound has a supremum or infimum respectively. Similarly, by transfer any internal subset of $^*\mathbb{N}$ must have a least element (i.e. a minimum) and if it is bounded from above, then it must have a greatest element (i.e. a maximum).

Suppose $\Phi(x)$ is internal and true for all $x \in \mathbb{N}$. Then $A = \{x \in {}^*\mathbb{N} \mid \neg\Phi(x)\}$ is an internal subset of $\mathbb{N}_\infty$, by Keisler's internal definition principle (1.3.7). Now, $A \subseteq {}^*\mathbb{N} \setminus \mathbb{N}$, by premise. Since, it is internal it must have a least element, as noted above. This least element must therefore be an element $x = \kappa \in \mathbb{N}_\infty$. Thus, $\Phi(x)$ is true up to some $\kappa - 1 \in \mathbb{N}_\infty$.

The two remaining cases follow similarly. □

Our next permanence principle applies to a broader context than Cauchy's principle (1.5.2). To state it we need the notion of a;

Definition 1.5.3 (directed set). A nonempty set $\mathcal{D}$ with a relation $\rhd$ is called a *directed set* if the relation $\rhd \subseteq \mathcal{D} \times \mathcal{D}$ is reflexive, transitive, and for each finite subset $U \in \mathcal{P}_F(\mathcal{D})$ there exists a $d \in \mathcal{D}$ such that $d \rhd d'$ for every $d' \in U$. In other words $\rhd$ is reflexive, transitive, and concurrent. If there is no element $m \in \mathcal{D}$ such that $m \rhd n$ for all $n \in \mathcal{D}$, we say that the directed set has no *terminal element*. Finally, if d, d' and $d \rhd d'$ we say that d is a successor of d', that d comes later than d', that d' is preceeds d, and that d' is earlier than d.

Given a directed set $(\mathcal{D}, \geq)$ the set of nearterminal elements $\mathcal{D}_\infty$ is defined

$$\mathcal{D}_\infty := \{d \in {}^*\mathcal{D} \mid \forall d' \in \mathcal{D}(d \rhd d')\}.$$

We call the following principle Overflow. However, other authors may refer to it as Overspill, Underflow, Spillover, etc.

Proposition 1.5.4 (Overflow)**.** *Suppose $(\mathcal{D}, \rhd)$ is a directed set with no terminal element, card($\mathcal{D}$) $< \kappa$, the transfer function $*$ is κ-saturated, and $\Phi(x)$ is an internal formula with a free variable x. Then the following hold:*

1. *If $\Phi(x)$ is true for all $x \in \mathcal{D}$, then $\Phi(x)$ is true for some $x \in \mathcal{D}_\infty$.*
2. *If $\Phi(x)$ is true for all $x \in \mathcal{D}_\infty$, then $\Phi(x)$ is true for some $x \in \mathcal{D}$.*

Proof. Let $(\mathcal{D}, \rhd)$, κ, and Φ be as in the statement of the proposition.

We begin by proving that the first assertion holds. Note first that $\rhd$ is $*$-concurrent, by transfer. In particular, $\rhd$ is internal and concurrent. Now, for each finite set $D \in \mathcal{P}_F(\mathcal{D})$ there exists a $d \in {}^*\mathcal{D}$ such that $\Phi(d)$ is true and $d \rhd d'$ for all $d' \in D$. For, there exists a $d \in \mathcal{D} \subseteq {}^*\mathcal{D}$ such that $d \rhd d'$ for all $d' \in D$, since $\geq$ is concurrent, and $\Phi(d)$ is true for all $d \in \mathcal{D}$. It follows that there is a $d \in {}^*D$ such that $\Phi(d)$ is true and $d \rhd d'$ for all $d' \in \mathcal{D}$, by κ-saturation (1.4.8). In other words, $\Phi(x)$ is true for some $x \in \mathcal{D}_\infty$.

The second assertion follows from the first, by the following argument. Suppose, $\Phi(x)$ is true for all $x \in \mathcal{D}_\infty$. Now, if $\Phi(x)$ isn't true for some $x \in \mathcal{D}$, then $\neg\Phi(x)$ is true for all $x \in \mathcal{D}$. It follows, that $\neg\Phi(x)$ is true for some $x \in \mathcal{D}_\infty$, by the first part of the proof. However, this contradicts our assumption that $\Phi(x)$ is true for all $x \in \mathcal{D}_\infty$. Thus, there must exist some $x \in \mathcal{D}$ such that $\Phi(x)$ holds. □

The reader should be warned that it is easy, at least initially, to make the mistake of applying something like Cauchy's principle to a formula which isn't in fact internal. In fact, this mistake, I wager, is by far the most common pitfall for anyone new to nonstandard analysis.

1.6 Basic Nonstandard Topology and Real Analysis

In this section we briefly outline the basic notions involved in nonstandard topology. We recall the following basic notions of topology;

Definition 1.6.1. A topology $\mathcal{T}$ on a set X is a collection of subsets of X containing both X and $\emptyset$, closed under finite intersection, and closed under arbitrary union. The complements of elements of $\mathcal{T}$ (with respect to X) are called *closed* sets. The tuple $(X, \mathcal{T})$ is called a *topological space*, although frequently we simply state X. A set $\mathcal{B} \subset \mathcal{T}$ is called a *basis* (or *base*) for the topology $\mathcal{T}$ if every element of $\mathcal{T}$ can be written as a union of elements of $\mathcal{B}$. That is, a family of open sets $\mathcal{B}$ is a basis for the space if every open set can be written as a union of its elements. For each $x \in X$ the *neighbourhood system of* x, denoted $\mathcal{T}_x$, is defined as the collection of subsets of X which are a superset of at least some open set containing x, i.e.

$$\mathcal{T}_x := \{U \subseteq X \mid \exists V \in \mathcal{T}(x \in V \wedge V \subseteq U)\}.$$

A set $\mathcal{B} \subset \mathcal{T}_x$ is called a *basis* (or *base*) for the neighbourhood system if for every element A in $\mathcal{T}_x$ there is an element $B \in \mathcal{B}$ such that $B \subset A$. Bases for neighbourhood systems are also referred to as *neighbourhood bases.*

A topological space $(X, \mathcal{T})$ is called *Hausdorff* if for every pair $(x, y) \in X^2$ of distinct points there are disjoint open sets U, V such that $x \in U$ and $y \in V$. Moreover a topological space $(X, \mathcal{T})$ is called *first countable* (or N_1) if every neighbourhood system has a countable (neighbourhood) basis.

A subset A of a space is called *compact* if for every open cover $\mathcal{C}$ of A – i.e. family of open sets $\mathcal{C}$ such that $A \subseteq \bigcup \mathcal{C}$ – there is a finite subset $\mathcal{D}$ of $\mathcal{C}$ which covers A. (A subset of a cover which also covers the set is called a *subcover*.)

In this section let X (or $(X, \mathcal{T})$) be a topological space.

Definition 1.6.2. The monad of a point $x \in X$ is the set

$$m(x) = \bigcap_U^{\mathcal{T}_x} {}^*U.$$

The points in $\bigcup_x^X m(x)$ are called the *nearstandard* points of ${}^*X$, denoted $\text{ns}(X)$. Moreover if $y \in m(x)$, then x is called the *standard part* of y. This is denoted ${}^\circ y = x$ or $\text{st}(y) = x$. If a point in ${}^*X$ is not nearstandard it is called *remote.*

Note that the standard part is not necessarily a well defined notion. Indeed it turns out that for it to be well defined we must restrict ourselves to Hausdorff spaces. But, before we prove this let us apply these defnitions to the following, frequently useful, propositions;

Proposition 1.6.3 (the basis monad is the monad)**.** *Let X be a topological space and $x \in X$. Then for every neighbourhood basis $\mathcal{B}$ of x we have the following equality*

$$\bigcap_b^{\mathcal{B}} {}^*b = m(x).$$

Proof. Let X be a topological space, $x \in X$, and $\mathcal{B}$ a neighbourhood basis for x. Since $\mathcal{B} \subset \mathcal{T}_x$ it follows that

$$\bigcap_b^{\mathcal{B}} {}^*b \supseteq \bigcap_n^{\mathcal{T}_x} {}^*n = m(x).$$

On the other hand, we can for each $n \in \mathcal{T}_x$ choose a $b_n \in \mathcal{B}$ such that $b_n \subset n$, by the definition of a neighbourhood basis, and, by transfer, ${}^*b_n \subset {}^*n$. Moreover these b_n form a neighbourhood basis for $\mathcal{T}_x$, again by definition. Now note that,

$$m(x) = \bigcap_{n}^{\mathcal{T}_x} {}^*n \supseteq \bigcap_{n}^{\mathcal{T}_x} {}^*b_n \supseteq \bigcap_{b}^{\mathcal{B}} {}^*b.$$

□

It turns out that we need a saturation argument to guarantee the existence of infinitesimal $*$-open sets, i.e. for any $x \in X$ there exists a $*$-open neighbourhood $U \in {}^*\mathcal{T} \cap {}^*\mathcal{T}_x$ such that $U \subset m(x)$. This is similar to the argument which shows that infinitessimal $*$-real numbers exist.

Proposition 1.6.4 (existence of infinitesimal $*$-open sets)**.** *Let X be a topological space and * be an enlargement. Then for every $x \in X$ there exists a $*$-open set $U \in {}^*\mathcal{T}$ such that $U \subset m(x)$.*

Proof. Let $\mathcal{B} = \mathcal{T}_x \cap \mathcal{T}$ be the set of open sets containing x, then there exists a $*$-finite collection $\mathcal{U}$ of $*$-open neighbourhoods of x such that $\mathcal{U} \supset \mathcal{B}$, by enlargement (1.4.2). By transfer, the $*$-open sets ${}^*\mathcal{T}$ are closed under $*$-finite intersection and so $U = \bigcap_u^{\mathcal{U}} u \in {}^*\mathcal{T} \cap {}^*\mathcal{T}_x$. Now note that, since ${}^\sigma\mathcal{B} \subset \mathcal{U}$

$$U = \bigcap_{u}^{\mathcal{U}} u \subseteq \bigcap_{b}^{{}^\sigma\mathcal{B}} b = m(x),$$

where the last equality is true since the basis monad is the monad (1.6.3). □

We now prove that the notion of the standard part for nearstandard elements of ${}^*X$ is well defined precisely when X is Hausdorff.

Proposition 1.6.5. *A space X is Hausdorff if and only if for every pair of distinct points $x, y \in X$, the monads $m(x)$ and $m(y)$ are disjoint.*

Proof. Suppose that the space X is Hausdorff and let $x, y \in X$ be distinct points. By definition, there must exist disjoint open sets U, V such that $x \in U$ and $y \in V$. Note that $U \in \mathcal{T}_x$ and $V \in \mathcal{T}_y$, i.e. U and V are in the neighbourhood systems of x and y, by definition. It follows that

$$m(x) = \bigcap_{N}^{\mathcal{T}_x} {}^*N \subseteq {}^*U \quad \text{and} \quad m(y) = \bigcap_{N}^{\mathcal{T}_y} N \subseteq {}^*V.$$

Moreover, ${}^*U \cap {}^*V = \emptyset$, by transfer, and therefore

$$m(x) \cap m(y) \subseteq {}^*U \cap {}^*V = \emptyset.$$

Conversely, suppose that $m(x) \cap m(y) = \emptyset$ for any pair of distinct points $x, y \in X$. Now, by the existence of infinitesimal $*$-open sets (1.6.4), there exists $*$-open infinitesimal neighbourhoods $U \in {}^*\mathcal{T}_x \cap {}^*\mathcal{T}$, $V \in {}^*\mathcal{T}_y \cap {}^*\mathcal{T}$ of x and y respectively. The sets U and V, must be disjoint since $U \subseteq m(x)$ and $V \subseteq m(y)$ and $m(x) \cap m(y) = \emptyset$. Formally we can therefore conclude that

$$\exists U, V \in {}^*\mathcal{T}(x \in U \wedge y \in V \wedge U \cap V = \emptyset).$$

It follows, by transfer, that the space is Hausdorff. □

The following relation is also common in nonstandard analysis;

Definition 1.6.6 (near equality)**.** We say that the points $x, y \in {}^*X$ are *(topologically) nearly equal* if

$$x \simeq y \Leftrightarrow \exists z \in X(x, y \in m(z)).$$

It is also read as x is nearly y.

Note that this relation only concerns itself with nearstandard points of ${}^*X$. Namely, the remote points of ${}^*X$ are precisely the points of ${}^*X$ which are not (topologically) nearly equal to any point of ${}^*X$.

Unless the topology is Hausdorff the definition for near equality given here is not necessarily ideal. Instead the definition,

$$x \simeq y \Leftrightarrow x \in m(y),$$

may be preferable in such contexts. (Here y is always a standard point.)

Moreover, in the context of metric spaces another notion of near equality is equally natural. Namely, two points in the nonstandard extension are (metrically) nearly equal if the distance between them is infinitesimal. This notion of near equality comes along with the metric monad,

$$\mu(x) = \bigcap_{r}^{\mathbb{R}_{>0}} {}^*B(x, r),$$

where $x \in {}^*X$. One difference between the topological monad and the metric monad is that the latter is defined for every point in ${}^*X$. While the topological monad is defined only for the standard points in ${}^*X$ (points in X.) The metric monad and the topological monad agree on the standard points. We mention the metric monad and near equality mostly for completeness. We will not need them in this work.

We will, however, need the following propositions which connect the standard notion of convergence with "near convergence" in the nonstandard setting.

Proposition 1.6.7 (Robinson's sequential lemma)**.** *Given a first countable Hausdorff space* $X, \mathcal{T}$ *if an internal* $*$*-sequence* $k : {}^*\mathbb{N} \to {}^*X$ *is such that* $k(n) \simeq x \in {}^*X$ *for every* $n \in \mathbb{N}$ *then* $k(n) \simeq x$ *up to some* $n = \kappa \in \mathbb{N}_\infty$*.*

It may at a glance appear as though one can simply apply Cauchy's principle (1.5.2) on $k(n) \simeq x$. However, this is not possible since $\simeq$ is an external notion.

Proof. Let $(X, \mathcal{T})$ and k be as in the proposition statement and let $\mathcal{T}_{{}^\circ x}$ be the neighbourhood system of the (standard) point ${}^\circ x$. Since X is first countable $\mathcal{T}_{{}^\circ x}$ has a countable neighbourhood basis $\mathcal{B}$. Now, choose a sequence $\eta : \mathbb{N} \to \mathcal{B}$ such that η is eventually in each $B \in \mathcal{B}$, i.e. for each $B \in \mathcal{B}$ there is an $n \in \mathbb{N}$ such that for all $m \in \mathbb{N}_{\geq n}$ we have $\eta(m) \subseteq B$. Note that it follows that $\eta[\mathbb{N}] = \{\eta(m) \mid m \in \mathbb{N}\}$ is a neighbourhood basis for x as well. Now extend η by transfer to ${}^*\eta : {}^*\mathbb{N} \to {}^*X$. Then the internal formula

$$k(n) \in \bigcap_{m}^{n} {}^*\eta(m)$$

holds for each $n \in \mathbb{N}$, and thus up to some $n = \kappa \in \mathbb{N}_\infty$, by Cauchy's principle (1.5.2). In particular, for all $\lambda \leq \kappa$ we have

$$k(n) \in \bigcap_{m}^{\lambda} \eta(m) = m(x).$$

For, the finite indices this is by assumption and for the infinite indices it holds since the basis monad is the monad (1.6.3). □

The following is essentially a generalisation of Robinson's sequential lemma and provides something like a "converse" result as well.

Proposition 1.6.8 (nonstandard convergence)**.** *Let* $(X, \mathcal{T})$ *be a first countable Hausdorff space and* $a \in X$ *be nearstandard. Then an internal* $*$*-sequence* $x \in {}^*\mathbb{N} \to {}^*X$ *has the property that* $x(n) \simeq a$ *for* n *in an initial segment of* $\mathbb{N}_\infty$ *precisely when* ${}^\circ x|_{\mathbb{N}} \to {}^\circ a$ *as* $n \in \mathbb{N} \to \infty$*.*

Recall that $^\circ x$ is the standard part of x.

Proof. Let $x : {}^*\mathbb{N} \to {}^*X$ be an internal $*$-sequence, $\mathcal{B}$ be a countable neigbourhood basis for $\mathcal{T}_{^\circ a}$, and $V : \mathbb{N} \to \mathcal{B}$ be a sequence such that $V_{n+1} \subseteq V_n$ and $x_n \in {}^*V_n$ for each $n \in \mathbb{N}$.

Extend V to an internal (in fact standard) $*$-sequence, ${}^*V$, by transfer. Now, for each $n \in \mathbb{N}$ we have

$$x_n \in \bigcap_m^{n+1} {}^*V_m,$$

and thus up to some $n = \kappa \in \mathbb{N}_\infty$, by Cauchy's principle (1.5.2). In particular,

$$x_n \in \bigcap_m^{n+1} {}^*V_m \subseteq \bigcap_m^{\mathbb{N}} {}^*V_m = \bigcap_B^{\mathcal{B}} {}^*B = m({}^\circ a),$$

since the the basis monad is the monad (1.6.3), for each $n \in \mathbb{N}_\infty$ such that $n \leq \kappa$.

Conversely, suppose that for each $V \in \mathcal{B}$ there is an $n \in \mathbb{N}$ such that $x(m) \in V$ for all $m \in \mathbb{N}_{\geq n}$. This means that the standard sequence $z = {}^\circ x|_{\mathbb{N}} : \mathbb{N} \to X$ has the property that

$$\forall V \in \mathcal{T}_{^\circ a} \exists n \in \mathbb{N} \forall m \in \mathbb{N}_{\geq n}(z(m) \in V)$$

and thus

$$\begin{aligned}
&\forall V \in \mathcal{T}_{^\circ a} \exists n \in \mathbb{N} \forall m \in \mathbb{N}_{\geq n}(z(m) \in V) \\
&\Leftrightarrow \forall V \in \mathcal{T}_{^\circ a} \exists n \in \mathbb{N} \forall m \in {}^*\mathbb{N}_{\geq n}({}^*z(m) \in {}^*V) \\
&\Rightarrow \forall V \in \mathcal{T}_{^\circ a} \forall m \in \mathbb{N}_\infty({}^*z(m) \in {}^*V) \\
&\Leftrightarrow \forall m \in \mathbb{N}_\infty({}^*z(m) \in \bigcap_{V \in \mathcal{T}_{^\circ a}} {}^*V) \\
&\Leftrightarrow \forall m \in \mathbb{N}_\infty({}^*z(m) \simeq a).
\end{aligned}$$

Now since $x(n) \simeq {}^*z(n)$ by definition, it follows that this is true up to some $n = \kappa \in \mathbb{N}_\infty$, by Robinson's sequential lemma (1.6.7). And since ${}^*z(n) \simeq a$ for all $n \in \mathbb{N}_\infty$ it follows, by transitivity of near equality, that $x(n) \simeq a$ up to κ. □

One important consequence of the above is that if an internal $*$-real $*$-sequence is infinite for all infinite indices then the standard part of the sequence grows without bound. Another consequence is that if a sequence $x : \mathbb{N} \to X$ converges to some point a then ${}^*x(n) \simeq a$ for all $n \in \mathbb{N}_\infty$ up to some $n = \kappa \in \mathbb{N}_\infty$. (In fact for standard sequences we have ${}^*x(n) \simeq a$ for all $\mathbb{N}_\infty$ precisely when x converges to a.) This also gives us a way to prove that not all $x : {}^*\mathbb{N} \to {}^*X$ are internal. For instance, let $x : \mathbb{N} \to X$ be any sequence which converges to a, then

$$x'(n) := \begin{cases} x(n), & n \in \mathbb{N}, \\ {}^*b, & n \in \mathbb{N}_\infty. \end{cases}$$

where $b \in X$ and $b \neq a$, cannot be internal, by nonstandard convergence (1.6.8).

For the next important nonstandard topological result we need the nonstandard characterisation of open sets.

Proposition 1.6.9 (nonstandard characterisation of open sets)**.** *Let X be a topological space. Then, a set A is open if and only if $m(x) \subseteq {}^*A$ for every point $x \in A$.*

Proof. Let A be an open set and $x \in A$. Then there is some open $V \in \mathcal{T}_x$ such that $V \subseteq A$, it follows, by transfer, that ${}^*V \subseteq {}^*A$. In particular,

$$m(x) \subseteq {}^*V \subseteq {}^*A.$$

Conversely, suppose A is a set in X such that $m(x) \subseteq {}^*A$ for each $x \in A$. Now, each $x \in A$ has an infinitesimal $*$-open neighbourhood B_x, by the existence of infinitesimal $*$-open sets (1.6.4). Then, $\forall x \in A \exists B \in {}^*\mathcal{T}_x(B \subseteq {}^*A)$ and thus $\forall x \in A \exists B \in \mathcal{T}_x(B \subseteq A)$, by transfer. Thus, A is open. □

We shall need the following nonstandard topological results later, when connecting the theory of Loeb measures to standard measure theory.

Theorem 1.6.10 (the standard image is closed)**.** *If* * *is* κ*-saturated and* $(X, \mathcal{T})$ *is a topological space such that card*$(\mathcal{T}) < \kappa$*, then for any internal* $U \subseteq X$ *the set*

$$st[U] = \{st(u) \mid u \in U\}$$

is closed in $\mathcal{T}$*.*

Proof. Let $x \notin \mathrm{st}[U]$. Then $m(x) \cap U = \emptyset$, by definition. Suppose now that there is no $V \in \mathcal{T}_x$ disjoint from $\mathrm{st}[U]$. This is equivalent to saying that each open neighbourhood V of x contains a point y such that $m(y) \cap U \neq \emptyset$. In particular, $V \cap U \supseteq m(y) \cap U \neq \emptyset$ for each $V \in {}^\sigma\mathcal{T}_x$, by the nonstandard characterisation of open sets (1.6.9). It follows that for any finite collection $V_i \in {}^\sigma\mathcal{T}_x$, $i \in n$ of $n \in \mathbb{N}$ sets the intersection $\bigcap_i^n V_i \in {}^\sigma\mathcal{T}_x$ and thus $\bigcap_i^n (V_i \cap U) \neq \emptyset$. In other words the collection $C = \{V \cap U \mid V \in {}^\sigma\mathcal{T}_x\}$ has the finite intersection property. Moreover, $\mathrm{card}(C) \leq \mathrm{card}(\mathcal{T}_x) \leq \mathrm{card}(\mathcal{T}) < \kappa$. It follows that $\bigcap C \neq \emptyset$, by the κ-intersection principle (1.4.10). But then,

$$\emptyset \neq \bigcap C = \bigcap_{V \in {}^\sigma\mathcal{T}_x} V \cap U = m(x) \cap U,$$

which contradicts our supposition. Thus, there must be a neighbourhood $V \in \mathcal{T}_x$ disjoint from $\mathrm{st}[U]$ and consequently the complement of $\mathrm{st}[U]$ is open. It follows that $\mathrm{st}[U]$ is closed, by definition. □

The following characterisation of compactness by Robinson, is, one of the most important nonstandard topological results.

Theorem 1.6.11 (Robinson's compactness criterion)**.** *A topological space* X *is compact if and only if* $ns(X) = {}^*X$*, that is all points in the nonstandard extension are nearstandard.*

Proof. Suppose that X is compact and that there is a remote point $y \in {}^*X$. Then, for each $x \in X$ we can choose an open set $B_x \in \mathcal{T}_x$ such that $y \notin {}^*B_x$. (In fact, $y \notin m(x) = \bigcap_b^{\mathcal{T}_x} {}^*b$, by definition.) This collection $\mathcal{B} = \{B_x | x \in X\}$ is clearly an open cover of X and there must therefore exist some finite subcover $\mathcal{C} = \{B_{x_i} | i \in n\} \subseteq \mathcal{B}$, by the compactness of X. Consequently ${}^*\mathcal{C} = \{{}^*B_{x_i} | i \in n\}$ is a cover of ${}^*X$, by transfer. But, this means $y \in {}^*B_{x_i}$ for some $x_i \in X$, $i \in n$, which is impossible by our initial choice criteria for B_x. It follows that if X is compact, then ${}^*X$ has no remote points. In other words every point of ${}^*X$ must be nearstandard.

Conversely, suppose that every point of ${}^*X$ is nearstandard and that X is not compact. Then there exists some open cover $\mathcal{C}$ of X with no finite subcover. It follows that the set ${}^*\mathcal{C}$ has no $*$-finite subcover, by transfer. However, by enlargement (1.4.2), there exists a $*$-finite subset $\mathcal{D} \subset {}^*\mathcal{C}$ such that ${}^\sigma\mathcal{C} \subset \mathcal{D}$. Moreover since every point is nearstandard ${}^\sigma\mathcal{C}$ covers ${}^*X$ and therefore so does $\mathcal{D}$. By transfer, it follows that $\mathcal{C}$ has a finite subcover. But, we chose $\mathcal{C}$ such that it had no finite subcover. It follows that no such $\mathcal{C}$ exists and consequently that X is compact. □

Finally, let us discuss some of the nonstandard topological resuts related to the real numbers. One elementary fact is that;

Proposition 1.6.12. *Every finite $*$-real number $r \in {}^*\mathbb{R}$ (that is every $r \in {}^*\mathbb{R}$ such that there exists an $s \in \mathbb{R}_{\geq 0}$ such that $|r| < s$) has a unique standard part. In particular, every finite $*$-real number is nearstandard.*

Proof. Let $r \in {}^*\mathbb{R}$ be finite. Consider the set $R = \{s \in \mathbb{R} \mid s \leq r\}$. Notice that R is non-empty, since it includes $-s$ for any $s \in \mathbb{R}$ such that $|r| < s$. Moreover, R is bounded from above by any $s \in \mathbb{R}$ such that $|r| < s$. It follows that $t = \sup R$ is well defined in $\mathbb{R}$. Suppose now that $|r - t| < \varepsilon$ for some $\varepsilon \in \mathbb{R}_{>0}$. Then either $t < r$ or $r < t$, by transfer. If $t < r$ then $t + \frac{\varepsilon}{2} < r$, contradicts that t is the supremum of R. (Indeed it follows that t isn't an upper bound of R at all.) On the other hand if $r < t$, then $r < t - \frac{\varepsilon}{2}$. This again contradicts the fact that t is the supremum of R, since $t - \frac{\varepsilon}{2}$ is a smaller upper bound of R than t. It follows that $|r - t| < \varepsilon$ for all $\varepsilon \in \mathbb{R}_{>0}$. But now, since $\mathcal{B} = \{B(t, \varepsilon) \mid \varepsilon \in \mathbb{R}_{>0}\}$ forms a neighbourhood basis for $\mathcal{T}_t$, we have

$$t \in \bigcap_{B}^{\mathcal{B}} {}^*B = m(t),$$

since the basis monad is the monad (1.6.3). □

We recall the definition of the extended real numbers and their topology;

Definition 1.6.13 ($\overline{\mathbb{R}}$)**.** Let $\overline{\mathbb{R}}$ denote the extended real line, i.e. $\mathbb{R} \cup \{-\infty, \infty\}$ along with the usual topology, i.e.

$$\mathcal{T}_{\overline{\mathbb{R}}} = \mathcal{T}_{\mathbb{R}} \cup \{[-\infty, a), (a, \infty] \mid a \in \mathbb{R}\}.$$

One notable feature of ${}^*\overline{\mathbb{R}}$ is that;

Proposition 1.6.14 ($\mathrm{ns}({}^*\overline{\mathbb{R}}) = {}^*\overline{\mathbb{R}}$)**.** *Every element of ${}^*\overline{\mathbb{R}}$ is nearstandard.*

Proof. The fact that the finite $*$-extended real numbers are nearstandard follows from Proposition 1.6.12.

It only remains to show that the infinite $*$-extended real numbers are nearstandard. Let κ be any positive infinite extended $*$-real $\kappa \in {}^*\overline{\mathbb{R}}_{\infty}$. Then,

$$\kappa \in \bigcap_{n}^{\mathbb{N}} {}^*(n, \infty] = m(\infty),$$

since the sets $(n, \infty]$, $n \in \mathbb{N}$ form a neighbourhood basis for $\infty \in \overline{\mathbb{R}}$ and the basis monad is the monad (1.6.3). Similarly, for any negative infinite extended $*$-real κ we have $\kappa \in m(-\infty)$. □

This naturally leads to the following observation;

Proposition 1.6.15 ($\overline{\mathbb{R}}$ is compact)**.** *The space $\overline{\mathbb{R}}$ is compact.*

Proof. Since every $*$-extended real number is nearstandard, $\mathrm{ns}({}^*\overline{\mathbb{R}}) = {}^*\overline{\mathbb{R}}$ (1.6.14), it follows that $\overline{\mathbb{R}}$ is compact, by Robinson's compactness criterion (1.6.11). □

Chapter 2

Loeb Measure Theory

Initially, early in nonstandard analysis, it was not quite clear how to fit measure theory into the framework productively. The breakthrough was made by Loeb in *Conversion from nonstandard to standard measure spaces and applications in probability theory* (Loeb 1975) which gave rise to the subject of Loeb measures/spaces.

We will present the theory of unbounded real valued Loeb measures. There are not many resources that treat the subject of the unbounded Loeb measure. The original paper Loeb (1975) is for bounded measures. Unbounded Loeb measures were first introduced by Henson. The most common references for them is *Foundations of Infinitesimal Stochastic Analysis* (Stroyan and Bayod 1986), for an elementary approach, and *Unbounded Loeb Measures* (Henson 1979) where they were first introduced. In practice, Loeb does also treat the subject of unbounded Loeb measures in *A functional approach to nonstandard measure theory* (Loeb 1984), where he presents a Daniell approach to the Loeb measure. (This approach is also presented in the book *An Introduction to Nonstandard Analysis* (Loeb and Hurd 1985).)

The theory of Loeb measures and the construction in particular has been extended to vector valued measures in *Nonstandard integration theory in topological vector lattices* (Loeb and Osswald 1997). However, we restrict ourselves to real valued measures.

Most applications are mainly concerned with bounded Loeb measures and probability measures in particular. In fact, the so called hyperfinite Loeb probability measures, which are extremely well behaved, turn out to be sufficient for essentially all purposes.

2.1 Loeb Measures

In this section we will construct the unbounded Loeb measure. Recall the following definitions from standard measure theory;

Definition 2.1.1 (algebra & σ-algebra)**.** A collection $\mathcal{A}$ is called an *algebra* on X if it satisfies the following criteria:

$$\mathcal{A} \subseteq \mathcal{P}(X),$$
$$X \in \mathcal{A},$$
$$A \in \mathcal{A} \Rightarrow A^{\mathsf{C}} = X \setminus A \in \mathcal{A},$$
$$A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}.$$

Moreover $\mathcal{A}$ is called a *σ-algebra* if it is closed under countable unions. That is for any $\{A_n\}_n^{\mathbb{N}} \subseteq \mathcal{A}$ it follows that $\bigcup_n^{\mathbb{N}} A_n \in \mathcal{A}$.

Definition 2.1.2 (measure)**.** A function $\mu : \mathcal{A} \to \mathbb{R}$ (or $\mathcal{A} \to \overline{\mathbb{R}}$) is called a real valued (extended real valued) *measure* if $\mathcal{A}$ is a σ-algebra, μ is non-negative, $\mu(\emptyset) = 0$, and μ is σ-additive, i.e. for any countable subset $\{A_n\}_n^{\mathbb{N}} \subset \mathcal{A}$ of pairwise disjoint sets we have

$$\mu\left(\bigcup_n^{\mathbb{N}} A_n\right) = \sum_n^{\mathbb{N}} \mu(A_n).$$

Definition 2.1.3 (charge)**.** A function $\mu : \mathcal{A} \to \mathbb{F}$, where $\mathbb{F}$ is a field, is called a $\mathbb{F}$-valued *charge* if $\mathcal{A}$ is an algebra, μ is non-negative, $\mu(\emptyset) = 0$, and μ is additive, i.e. for any disjoint $A, B \in \mathcal{A}$ we have that $\mu(A \cup B) = \mu(A) + \mu(B)$.

What we call a charge here is also known as a *finitely additve measure*. We prefer the term 'charge' as it reduces the risk of conflating the two.

It is easily shown that any algebra (and σ-algebra) $\mathcal{A}$ contains the empty set and is closed under difference and intersection, i.e. $\emptyset \in \mathcal{A}$ and for each $A, B \in \mathcal{A}$ we have that $A \cap B \in \mathcal{A}$ and $A \setminus B \in \mathcal{A}$. Simply note that $\emptyset = X \setminus X \in \mathcal{A}$ and

$$A \cap B = (A^{\mathsf{C}} \cup B^{\mathsf{C}})^{\mathsf{C}} \in \mathcal{A} \quad \text{and} \quad A \setminus B = A \cap B^{\mathsf{C}} \in \mathcal{A}.$$

Furthermore, they are also closed under finite union and intersection, as can be shown by a simple induction argument.

By equally simple considerations it is clear that both charges and measures are finitely additive;

Definition 2.1.4 (finitely additive)**.** A set function $\mu : \mathcal{A} \to \overline{\mathbb{R}}$ is called *finitely additive* if for each finite collection $\{A_n\}_n^m \subseteq \mathcal{A}$ of $m \in \mathbb{N}$ pairwise disjoint sets we have that

$$\mu\left(\bigcup_n^{m} A_n\right) = \sum_n^{m} \mu(A_n).$$

and monotone;

Definition 2.1.5 (monotone)**.** A set function $\mu : \mathcal{A} \to \overline{\mathbb{R}}$ is called *monotone* if for each $A, B \in \mathcal{A}$ such that $A \subseteq B$ we have that $\mu(A) \leq \mu(B)$.

Simple arguments also reveal that a charge (or measure) is additive if and only if it is finitely additive. From now on we will take these properties for granted.

The Loeb measure construction starts from an internal ${}^*\mathbb{R}$-valued charge on an internal set algebra. Importantly, these objects have the following properties;

Proposition 2.1.6. *Internal set algebras are closed under $*$-finite intersection and union, i.e. if $\mathcal{A}$ is an internal algebra and $\{A_n \mid n \in \kappa\} \subseteq \mathcal{A}$, where $\kappa \in {}^*\mathbb{N}$, then*

$$\bigcap_n^{\kappa} A_n \in \mathcal{A} \quad \text{and} \quad \bigcup_n^{\kappa} A_n \in \mathcal{A}.$$

Proof. Let $\mathcal{A}$ be an internal algebra on an internal set X. The proof is by $*$-induction (1.3.11). Clearly, for $n = 0$,

$$\bigcap_m^0 A_m = X, \bigcup_m^0 A_m = \emptyset,$$

where we use the convention that the empty intersection is the full set and the empty union is the empty set. Let $n \in {}^*\mathbb{N}$ and $\{A_m \mid m \in (n+1)\} \subseteq \mathcal{A}$, Suppose now that the proposition holds up to $n \in {}^*\mathbb{N}$, i.e.

$$B_n \cup \bigcup_m^n A_n \in \mathcal{A},$$

then

$$\bigcup_m^{n+1} A_m = A_n \cup \bigcup_m^n A_m = A_n \cup B_n \in \mathcal{A}$$

and similarly for intersection. Thus, the proposition follows by $*$-induction (1.3.11). □

Proposition 2.1.7. *Internal charges are $*$-finitely additive, i.e. for any internal charge $\mu : \mathcal{A} \to {}^*\mathbb{R}$ if $\{A_n \mid n \in \kappa\} \subseteq \mathcal{A}$ of $\kappa \in {}^*\mathbb{N}$ pairwise disjoint sets, then*

$$\mu\left(\bigcup_n^{\kappa} A_n\right) = \sum_n^{\kappa} \mu(A_n).$$

Proof. Let $\mathcal{A}$ be an internal algebra and $\mu : \mathcal{A} \to {}^*\mathbb{R}$ be an internal charge. The proof is by $*$-induction (1.3.11). Clearly, for $n = 0$,

$$\mu\left(\bigcup_m^0 A_m\right) = \mu(\emptyset) = 0,$$

where we have again used the convention that the empty union is the empty s et. Let $n \in {}^*\mathbb{N}$ and $\{A_m \mid m \in (n+1)\} \subseteq \mathcal{A}$ be a set of $n+1$ pairwise disjoint sets in $\mathcal{A}$. Suppose now that the the proposition holds up to $n \in {}^*\mathbb{N}$, i.e.

$$\mu\left(\bigcup_m^n A_m\right) = \sum_m^n \mu(A_n),$$

then

$$\begin{aligned}
\mu\left(\bigcup_m^{n+1} A_m\right) &= \mu\left(A_n \cup \bigcup_m^n A_m\right) \\
&= \mu(A_n) + \mu\left(\bigcup_m^n A_m\right) \\
&= \mu(A_n) + \sum_m^n \mu(A_m) \\
&= \sum_m^{n+1} \mu(A_m).
\end{aligned}$$

Thus, the proposition follows by $*$-induction (1.3.11). □

From this point on, let X be some internal set, $\mathcal{A}$ be an internal algebra on X, and $\mu : \mathcal{A} \to {}^*\mathbb{R}$ be an internal charge. This function μ will be referred to as the *charge* of the Loeb measure. It is also worth noting that μ is always $*$-bounded, i.e. there exists an $r \in {}^*\mathbb{R}$ such that $\mu(A) \leq r$ for all $A \in \mathcal{A}$, since $\mu(X) \in {}^*\mathbb{R}$, by definition, and thus $\mu(A) \leq \mu(X)$ for all $A \in \mathcal{A}$. Moreover, when we say μ is unbounded, we mean that the range of μ is not bounded by any $r \in \mathbb{R}$ (read $r \in {}^\sigma\mathbb{R}$).

We want to "extend" this μ, or rather ${}^\circ\mu$, into an extended real valued σ-additive measure. Since $^\circ$ is a morphism of addition and any internal algebra is an algebra, it follows that ${}^\circ\mu$ has many the properties of a measure. (Recall that $^\circ : {}^*\mathbb{R} \to \mathbb{R}$ is the standard part function.) In particular it is a charge. However, we want to extend $\mathcal{A}$ to a σ-algebra on which ${}^\circ\mu$ is σ-additive. The σ-algebra $\sigma(\mathcal{A})$ induced by $\mathcal{A}$ is a possible candidate. However, we will build an even larger σ-algebra. To do this we will use the Caratheodory extension of the outer measure. For which we need.

Definition 2.1.8 (standard part outer measure)**.** The outer measure induced by μ is defined and denoted

$$\overline{\mu}(U) = \inf\{{}^\circ\mu(A) \mid U \subseteq A \in \mathcal{A}\},$$

for each $U \subseteq X$.

Note that in the previous definition and all following definitions we use the general subset relation rather than the internal one. This means that the subsets may be external sets.

We begin by proving that the standard part outer measure (2.1.8) $\overline{\mu}$ is an outer measure.

Definition 2.1.9 (outer measure)**.** A function $m : \mathcal{P}(X) \to [0, \infty]$ is called an *outer measure* if it assigns $\emptyset$ to 0, it is monotone, and σ-subadditive. In other words m has the following properties:

$$\begin{aligned}
&m(\emptyset) = 0, \\
&A \subseteq B \Rightarrow m(A) \leq m(B) && \forall A, B \subseteq X, \\
&m\left(\bigcup_n^{\mathbb{N}} A_n\right) \leq \sum_n^{\mathbb{N}} m(A_n), && \forall A \in \mathcal{P}(X)^{\mathbb{N}}.
\end{aligned}$$

Proposition 2.1.10. *The standard part outer measure $\overline{\mu}$ of μ is an outer measure on X.*

Proof. The monotone property is a direct consequence of the properties of inf. If $A \subseteq B \subseteq X$ then,

$$\{{}^\circ\mu(C) \mid A \subseteq C \subseteq X\} \supset \{{}^\circ\mu(C) \mid B \subseteq C \subseteq X\} \tag{2.1}$$

and thus

$$\overline{\mu}(A) = \inf\{{}^\circ\mu(C) \mid A \subseteq C \subseteq X\} \leq \inf\{{}^\circ\mu(C) \mid B \subseteq C \subseteq X\} = \overline{\mu}(B). \tag{2.2}$$

Now, let $A_n \subseteq X$ for each $n \in \mathbb{N}$. If $\overline{\mu}(A_m) = \infty$ for some $m \in \mathbb{N}$, then $\overline{\mu}\left(\bigcup_n^{\mathbb{N}} A_n\right) = \infty$, since $\overline{\mu}$ is monotone, and thus

$$\overline{\mu}\left(\bigcup_n^{\mathbb{N}} A_n\right) \leq \sum_n^{\mathbb{N}} \overline{\mu}(A_n).$$

On the other hand suppose $\overline{\mu}(A_n) < \infty$ for every $n \in \mathbb{N}$. Then for each $\varepsilon \in \mathbb{R}_{>0}$ we can choose a sequence $B_n \in \mathcal{A}$, $n \in \mathbb{N}$, such that $A_n \subseteq B_n$ and $\overline{\mu}(B_n) \leq \overline{\mu}(A_n) + \frac{\varepsilon}{2^n}$. We can extend both $\overline{\mu}(A_n)$ and B_n to internal $*$-sequences, by the κ-extension principle (1.4.11). Since

$$\mu(B_n) \leq \overline{\mu}(A_n) + \frac{\varepsilon}{2^n},$$

is true for each $n \in \mathbb{N}$ it follows that it holds up to some $n = N \in \mathbb{N}_\infty$, by Cauchy's principle (1.5.2). Then for each $\kappa \in \mathbb{N}_\infty$ such that $\kappa \leq N$ we have that

$$\overline{\mu}\left(\bigcup_n^{\mathbb{N}} A_n\right) \leq \overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right) \leq \overline{\mu}\left(\bigcup_n^{\kappa} B_n\right) = {}^\circ\mu\left(\bigcup_n^{\kappa} B_n\right) \leq {}^\circ\sum_n^{\kappa} \mu\left(B_n\right)$$
$$\leq {}^\circ\sum_n^{\kappa}\left(\overline{\mu}(A_n) + \frac{\varepsilon}{2^n}\right) \leq {}^\circ\left(\sum_n^{\kappa}\overline{\mu}(A_n)\right) + 2\varepsilon.$$

It follows that $\overline{\mu}\left(\bigcup_n^{\mathbb{N}} A_n\right) \leq {}^\circ\left(\sum_n^{\kappa}\overline{\mu}(A_n)\right)$ for every $\kappa \in \mathbb{N}_\infty$ such that $\kappa \leq N$. Moreover, since we know that the limit $\sum_n^{\mathbb{N}}\overline{\mu}(A_n)$ exists (in $[-\infty, \infty]$) it follows that

$$\sum_n^{\lambda}\overline{\mu}(A_n) \simeq \sum_n^{\mathbb{N}}\overline{\mu}(A_n),$$

for all λ in some initial segment of $\mathbb{N}_\infty$, by nonstandard convergence (1.6.8). Thus

$$\overline{\mu}\left(\bigcup_n^{\mathbb{N}} A_n\right) \leq {}^\circ\sum_n^{\kappa}\overline{\mu}(A_n) = \sum_n^{\mathbb{N}}\overline{\mu}(A_n),$$

for a sufficiently small $\kappa \in \mathbb{N}_\infty$. □

At this point one can employ the Caratheodory extension theorem to get the σ-algebra and measure related to the standard part outer measure. The σ-algebra is the Loeb algebra and the measure is the Loeb measure. See for instance *Foundations of Modern Analysis* (Friedman 2010).

We will instead opt for a presentation more closely aligned with Stroyan and Bayod 1986. The two are equivalent. We choose the latter only because it introduces some intermediate results and nomenclature which we shall find useful later on. The reader will also find results not included here and a longer discussion with regards to the unbounded Loeb measure in the aforementioned work by Stroyan and Bayod.

Definition 2.1.11. A subset $A \subseteq X$ is called

- *μ-measurable* if $A \in \mathcal{A}$,
- *$\overline{\mu}$-finite* if $\overline{\mu}(A) < \infty$ (alternatively when $A \in \mathcal{A}$ we can say A is μ-finite),
- *μ-approximable* if there exists a set $B \in \mathcal{A}$ such that $\overline{\mu}(A\Delta B) = 0$,
- *$\overline{\mu}$-integrable* if it is μ-approximable and μ-finite,
- and *$\overline{\mu}$-measurable* if the intersection $A \cap B$ is $\overline{\mu}$-integrable for every $\overline{\mu}$-integrable B. A subset $A \subseteq X$ is called *μ-integrable* if it is μ-approximable and it has finite outer measure.

The collection of μ-approximable sets is denoted $\Delta(\mu)$, the collection of $\overline{\mu}$-integrable sets is denoted $\mathcal{I}(\mu)$, and the collection of $\overline{\mu}$-measurable sets is denoted $\mathcal{M}(\mu)$.

Here $A\Delta B$ for sets A, B denotes the symmetric difference, that is

$$A\Delta B := (A \setminus B) \cup (B \setminus A).$$

From the definitions of the standard part outer measure (2.1.8) and of the μ-approximable sets it follows directly that;

Proposition 2.1.12 (ε-characterisation of $\Delta(\mu)$)**.** *A subset $A \subseteq X$ is a μ-approximable set, i.e. there is a $B \in \mathcal{A}$ such that $\overline{\mu}(A\Delta B) = 0$, if and only if for each $\varepsilon \in \mathbb{R}_{>0}$ there is a $C \in \mathcal{A}$ such that $A\Delta B \subseteq C$ and $\mu(C) < \varepsilon$.*

Proposition 2.1.13 (flanking characterisation of $\Delta(\mu)$)**.** *A subset $B \subseteq X$ is μ-approximable if and only if for each $\varepsilon \in \mathbb{R}_{>0}$ there are μ-measurable sets $A, C \in \mathcal{A}$ such that $A \subseteq B \subseteq C$ and $\mu(C \setminus A) < \varepsilon$.*

Proof. Suppose $B \subseteq X$ is μ-approximable. Then, for each $\varepsilon \in \mathbb{R}_{>0}$, there are sets $D, E \in \mathcal{A}$ such that $\overline{\mu}(B\Delta D) = 0$, $B\Delta D \subseteq E$, and $\overline{\mu}(E) < \varepsilon$, by the ε-characterisation of $\Delta(\mu)$ (2.1.12). Let $A = D \setminus E$ and $B = D \cup E$. Now $A, B \in \mathcal{A}$ since $\mathcal{A}$ is an algebra and $A \subseteq B \subseteq C$, since

$$A = D \setminus E \subseteq D \setminus (B\Delta D) = D \cap B \subseteq B$$

and

$$C = D \cup E \supseteq D \cup (B\Delta D) \supseteq B.$$

Finally,

$$\overline{\mu}(C \setminus A) = \overline{\mu}((D \cup E) \setminus (D \setminus E)) = \overline{\mu}(E) < \varepsilon.$$

□

In the following propositions we build up the Loeb measure.

Proposition 2.1.14. *The countable union of a of μ-measurable sets is $\overline{\mu}$-integrable if the union is $\overline{\mu}$-finite.*

Proof. Let $B \in \mathcal{A}^{\mathbb{N}}$ and $\overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right) < \infty$. We extend B to a internal $*$-sequence using the κ-extension principle (1.4.11). Then,

$$\overline{\mu}\left(\bigcup_n^m B_n\right) \leq \overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right) \leq \overline{\mu}\left(\bigcup_n^{\lambda} B_n\right),$$

for every $m \in \mathbb{N}$ and $\lambda \in \mathbb{N}_\infty$. Moreover,

$$\mu\left(\bigcup_n^m B_n\right) - \frac{1}{m} \leq \overline{\mu}\left(\bigcup_n^m B_n\right) \leq \overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right),$$

is true for all $m \in \mathbb{N}$ and thus the internal proposition

$$\mu\left(\bigcup_n^m B_n\right) - \frac{1}{m} \leq \overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right),$$

is true up to some $m = \kappa \ \in \mathbb{N}_\infty$, by Cauchy's principle (1.5.2). Now,

$$\overline{\mu}\left(\bigcup_n^{\lambda} B_n\right) \simeq \mu\left(\bigcup_n^{\lambda} B_n\right) - \frac{1}{\lambda} \leq \overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right) \leq \overline{\mu}\left(\bigcup_n^{\lambda} B_n\right),$$

for all $\lambda \in \mathbb{N}_\infty$ such that $\lambda \leq \kappa$. In other words

$$\mu\left(\bigcup_n^{\lambda} B_n\right) \simeq \overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right),$$

for all λ on an initial segment of $\mathbb{N}_\infty$. Therefore,

$$\overline{\mu}\left(\bigcup_n^m B_n\right) = {}^\circ\mu\left(\bigcup_n^m B_n\right) \to \overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right),$$

as $m \to \infty$ ($m \in \mathbb{N}$), by nonstandard convergence (1.6.8).

We now show that $\bigcup_n^{\mathbb{N}} B_n$ is μ-approximable. Observe that for any $\varepsilon \in \mathbb{R}_{>0}$ there is an $m \in \mathbb{N}$ such that

$$\begin{aligned}
\overline{\mu}\left(\left(\bigcup_n^\kappa B_n\right) \Delta \left(\bigcup_n^{\mathbb{N}} B_n\right)\right) &= \overline{\mu}\left(\left(\bigcup_n^\kappa B_n\right) \setminus \left(\bigcup_n^{\mathbb{N}} B_n\right)\right) \\
&\leq \overline{\mu}\left(\left(\bigcup_n^\kappa B_n\right) \setminus \left(\bigcup_n^m B_n\right)\right) < \mu\left(\left(\bigcup_n^\kappa B_n\right) \setminus \left(\bigcup_n^m B_n\right)\right) + \frac{1}{m} \\
&= \mu\left(\bigcup_n^\kappa B_n\right) - \mu\left(\bigcup_n^m B_n\right) + \frac{1}{m} < \overline{\mu}\left(\bigcup_n^\kappa B_n\right) - \overline{\mu}\left(\bigcup_n^m B_n\right) + \frac{2}{m} \\
&= \overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right) - \overline{\mu}\left(\bigcup_n^m B_n\right) + \frac{2}{m} < \varepsilon.
\end{aligned}$$

Note that here we rely on the fact that $\overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right) < \infty$ twice. The first time is at the strict inequality. Which is only true if $\overline{\mu}\left(\left(\bigcup_n^\lambda B_n\right) \setminus \left(\bigcup_n^{\mathbb{N}} B_n\right)\right) < \infty$. The second time is at the step where ε is introduced. For, it is true that given an $\varepsilon \in \mathbb{R}_{>0}$ the inequality

$$\left|\overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right) - \overline{\mu}\left(\bigcup_n^m B_n\right)\right| < \varepsilon$$

is satisfied for sufficiently large $m \in \mathbb{N}$ (assuming $\lim_{m\to\infty} \overline{\mu}\left(\bigcup_n^m B_n\right) = \overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right)$) in general only when $\overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right) < \infty$.

It follows that $\bigcup_n^{\mathbb{N}} B_n$ is $\overline{\mu}$-integrable by definition, since we have shown it to be μ-approximable and we assumed it is $\overline{\mu}$-finite. ☐

Proposition 2.1.15 (triangle inequality for sets)**.** *For any sets A, B, C the following set inclusions hold*

$$A \setminus C \subseteq (A \setminus B) \cup (B \setminus C) \text{ and } A\Delta C \subseteq (A\Delta B) \cup (B\Delta C).$$

Proof. The second follows from the first. The proof is left to the reader. ☐

Proposition 2.1.16. *The $\overline{\mu}$-integrable sets are closed under countable intersection, difference, and dominated countable union. In other words, for any $A \in \mathcal{I}(\mu)^{\mathbb{N}}$ we have that $\bigcap_n^{\mathbb{N}} A_n \in \mathcal{I}(\mu)$, for any pair $A, B \in \mathcal{I}(\mu)$ we have that $A \setminus B \in \mathcal{I}(\mu)$, and for any $A \in \mathcal{I}(\mu)^{\mathbb{N}}$, such that there is a $B \in \mathcal{I}(\mu)$ satisfying $\bigcup_n^{\mathbb{N}} A_n \subseteq B$, we have that $\bigcup_n^{\mathbb{N}} A_n \in \mathcal{I}(\mu)$.*

Proof. We begin by showing that the $\overline{\mu}$-integrable sets ($\mathcal{I}(\mu)$) are closed under set difference. Suppose $F, G \in \mathcal{I}(\mu)$. Then by definition there are sets $H, K \in \mathcal{A}$ such that

$$\overline{\mu}(F\Delta H) = \overline{\mu}(G\Delta K) = 0.$$

Now note that $\overline{\mu}(F \setminus G) \leq \overline{\mu}(F) < \infty$, $H \setminus K \in \mathcal{A}$, and that

$$\begin{aligned}
\overline{\mu}((F \setminus G)\Delta(H \setminus K)) &= \overline{\mu}(((F \setminus G) \setminus (H \setminus K)) \cup ((H \setminus K) \setminus (F \setminus G))) \\
&= \overline{\mu}((F \cap G^\complement \cap (H^\complement \cup K)) \cup (H \cap K^\complement \cap (F^\complement \cup G))) \\
&\leq \overline{\mu}((G^\complement \cap K) \cup (H \cap F^\complement)) \\
&\leq \overline{\mu}(G^\complement \cap K) + \overline{\mu}(H \cap F^\complement) \\
&\leq \overline{\mu}(G\Delta K) + \overline{\mu}(F\Delta H) = 0.
\end{aligned}$$

Thus $F \setminus G \in \mathcal{I}(\mu)$ by definition.

We now show that the $\overline{\mu}$-integrable sets are closed under dominated countable union. Suppose $A \in \mathcal{I}(\mu)^{\mathbb{N}}$ and $B \in \mathcal{I}(\mu)$ such that $\bigcup_n^{\mathbb{N}} A_n \subseteq B$. Note that $\bigcup_n^{\mathbb{N}} A_n$ is $\overline{\mu}$-finite, since $\overline{\mu}\left(\bigcup_n^{\mathbb{N}} A_n\right) \leq \overline{\mu}(B) < \infty$. Furthermore, there is a $B_n \in \mathcal{A}$ for each $n \in \mathbb{N}$ such that $\overline{\mu}(A_n \Delta B_n) = 0$, since each A_n is μ-approximable. Now observe that,

$$\begin{aligned}
\overline{\mu}\left(\left(\bigcup_n^{\mathbb{N}} A_n\right) \Delta \left(\bigcup_n^{\mathbb{N}} B_n\right)\right) &= \overline{\mu}\left(\left(\bigcup_n^{\mathbb{N}} A_n \setminus \bigcup_m^{\mathbb{N}} B_m\right) \cup \left(\bigcup_n^{\mathbb{N}} B_n \setminus \bigcup_m^{\mathbb{N}} A_m\right)\right) \\
&\leq \overline{\mu}\left(\left(\bigcup_n^{\mathbb{N}} (A_n \setminus B_n)\right) \cup \left(\bigcup_n^{\mathbb{N}} (B_n \setminus A_n)\right)\right) = \overline{\mu}\left(\bigcup_n^{\mathbb{N}} ((A_n \setminus B_n) \cup (B_n \setminus A_n))\right) \\
&= \overline{\mu}\left(\bigcup_n^{\mathbb{N}} (A_n \Delta B_n)\right) \leq \sum_n^{\mathbb{N}} \overline{\mu}(A_n \Delta B_n) = 0.
\end{aligned}$$

It follows that $\overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right) = \overline{\mu}\left(\bigcup_n^{\mathbb{N}} A_n\right) < \infty$. For, let $A' = \bigcup_n^{\mathbb{N}} A_n$ and $B' = \bigcup_n^{\mathbb{N}} B_n$ and observe that

$$\begin{aligned}
\overline{\mu}(A' \cup B') &\leq \overline{\mu}(A' \Delta B') + \overline{\mu}(A' \cap B') \\
&\leq \overline{\mu}(A') \quad (\text{or } \overline{\mu}(B')) \\
&\leq \overline{\mu}(A' \cup B').
\end{aligned}$$

Consequently, $\bigcup_n^{\mathbb{N}} B_n \in \mathcal{I}(\mu)$, by Proposition 2.1.14, and thus there exists a set $C \in \mathcal{A}$ such that $\overline{\mu}\left(C \Delta \bigcup_n^{\mathbb{N}} B_n\right)$. Finally note that $\bigcup_n^{\mathbb{N}} A_n$ is μ-approximable, since

$$\begin{aligned}
\overline{\mu}\left(C \Delta \bigcup_n^{\mathbb{N}} A_n\right) &\leq \overline{\mu}\left(\left(C \Delta \bigcup_n^{\mathbb{N}} B_n\right) \cup \left(\bigcup_n^{\mathbb{N}} B_n \Delta \bigcup_n^{\mathbb{N}} A_n\right)\right) \\
&\leq \overline{\mu}\left(C \Delta \left(\bigcup_n^{\mathbb{N}} B_n\right)\right) + \overline{\mu}\left(\left(\bigcup_n^{\mathbb{N}} B_n\right) \Delta \left(\bigcup_n^{\mathbb{N}} A_n\right)\right) \\
&= 0.
\end{aligned}$$

(Here we used the triangle inequality for sets (2.1.15) for the first inequality.) Which proves that $\bigcup_n^{\mathbb{N}} A_n \in \mathcal{I}(\mu)$, since we have shown it is both μ-approximable and $\overline{\mu}$-finite.

Finally, we show that the $\overline{\mu}$-integrable sets are closed under countable intersection. Suppose $A \in \mathcal{I}(\mu)^{\mathbb{N}}$. Then $\bigcap_n^{\mathbb{N}} A_n \in \mathcal{I}(\mu)$ follows from the previous two properties we have shown, since

$$\begin{aligned}
\bigcap_n^{\mathbb{N}} A_n = A_0 \cap \bigcap_n^{\mathbb{N}} A_n = A_0 \cap \left(\bigcap_n^{\mathbb{N}} A_n\right)^{\mathsf{CC}} &= A_0 \cap \left(\bigcup_n^{\mathbb{N}} A_n^{\mathsf{C}}\right)^{\mathsf{C}} = A_0 \setminus \bigcup_n^{\mathbb{N}} A_n^{\mathsf{C}} \\
&= A_0 \setminus \bigcup_n^{\mathbb{N}} (X \setminus A_n) = A_0 \setminus \bigcup_n^{\mathbb{N}} (A_0 \setminus A_n).
\end{aligned}$$

□

As a consequence we can conclude that;

Corollary 2.1.17 ($\mathcal{I}(\mu) \subseteq \mathcal{M}(\mu)$)**.** *All $\overline{\mu}$-integrable sets are $\overline{\mu}$-measurable.*

Proof. If $A \in \mathcal{I}(\mu)$ then for any $B \in \mathcal{I}(\mu)$ we have $A \cap B \in \mathcal{I}(\mu)$, by Proposition 2.1.16. Thus $A \in \mathcal{M}(\mu)$. □

In fact, more is true.

Proposition 2.1.18. *The $\overline{\mu}$-finite $\overline{\mu}$-measurable sets are precisely the $\overline{\mu}$-integrable sets.*

Proof. Suppose A is a $\overline{\mu}$-finite and $\overline{\mu}$-measurable. Then for any given $\varepsilon \in \mathbb{R}_{>0}$ there is a μ-measurable set $B \in \mathcal{A}$ such that $A \subseteq B$ and $\mu(B) \leq \overline{\mu}(A) + \varepsilon$. Notably, B is clearly $\overline{\mu}$-finite and μ-approximable. Thus B is $\overline{\mu}$-integrable. It follows that $A = A \cap B \in \mathcal{I}(\mu)$, by Proposition 2.1.16.

The converse is the above corollary that $\mathcal{I}(\mu) \subseteq \mathcal{M}(\mu)$ (2.1.17). □

We can now prove that;

Proposition 2.1.19. *The $\overline{\mu}$-measurable sets $\mathcal{M}(\mu)$ form a σ-algebra on X.*

Proof. Firstly we note that $X \in \mathcal{M}(\mu)$.

Secondly, we show that $\mathcal{M}(\mu)$ is closed under complements (relative to X). Suppose $A \in \mathcal{M}(\mu)$ and $B \in \mathcal{I}(\mu)$. Then $(X \setminus A) \cap B = B \setminus (A \cap B) \in \mathcal{I}(\mu)$. Thus $X \setminus A \in \mathcal{M}(\mu)$.

Thirdly, we show that $\mathcal{M}(\mu)$ is closed under countable intersections. Suppose $A_n \in \mathcal{M}(\mu)$ for each $n \in \mathbb{N}$ and $B \in \mathcal{I}(\mu)$. Then

$$\left(\bigcap_n^{\mathbb{N}} A_n\right) \cap B = \bigcap_n^{\mathbb{N}} (A_n \cap B) \in \mathcal{I}(\mu),$$

by Proposition 2.1.16. Thus $\bigcap_n^{\mathbb{N}} A_n \in \mathcal{M}(\mu)$.

Finally, we show that $\mathcal{M}(\mu)$ is closed under countable unions. Suppose $A_n \in \mathcal{M}(\mu)$ for each $n \in \mathbb{N}$. Then $\bigcup_n^{\mathbb{N}} A_n \in \mathcal{M}(\mu)$, since

$$\bigcup_n^{\mathbb{N}} A_n = \left(\bigcap_n^{\mathbb{N}} A^{\complement}\right)^{\complement} \in \mathcal{M}(\mu),$$

by the previous parts of the proof. □

From this point on we will refer to the σ-algebra $\mathcal{M}(\mu)$ as the Loeb algebra of μ and denote it $\mathcal{A}_L$. It is with respect to this σ-algebra that the standard part outer measure is a measure.

Proposition 2.1.20. *The standard part outer measure $\overline{\mu}$ restricted to the Loeb algebra is a measure.*

Proof. Since $\overline{\mu}$ is an outer measure it suffices to show that it is σ-additive on the Loeb algebra $\mathcal{A}_L$.

Let $A_n \in \mathcal{M}(\mu)$, $n \in \mathbb{N}$ be a sequence of disjoint sets. In the case $\overline{\mu}(A_m) = \infty$ for some $m \in \mathbb{N}$, we have that

$$\infty = \overline{\mu}(A_m) \leq \overline{\mu}\left(\bigcup_n A_n\right) \leq \sum_n^{\mathbb{N}} \overline{\mu}(A_n).$$

Thus $\overline{\mu}\left(\bigcup_n^{\mathbb{N}} A_n\right) = \sum_n^{\mathbb{N}} \overline{\mu}(A_n)$.

Suppose now that each $\overline{\mu}(A_n) < \infty$ for every $n \in \mathbb{N}$. Then each A_n, $n \in \mathbb{N}$ is $\overline{\mu}$-integrable and thus $\overline{\mu}$-approximable, by Proposition 2.1.18. Then for each A_n we can choose $B_n \in \mathcal{A}$ such that $\overline{\mu}(B_n \Delta A_n) = 0$. We can choose B_n such that it is pairwise disjoint, since for any B_n we can switch to a pairwise disjoint family B'_n by the following standard procedure. Let $B'_0 = B_0$ and $B'_n = B_n \setminus \bigcup_m^n B_m$. Then $B'_n \in \mathcal{A}$ for each $n \in \mathbb{N}$ and is the sequence is disjoint.

Furthermore,

$$\begin{aligned}
\overline{\mu}(A_n \Delta B_n') &= \overline{\mu}\left(A_n \Delta \left(B_n \setminus \bigcup_m^n B_m\right)\right) \\
&= \overline{\mu}\left(\left(A_n \setminus \left(B_n \setminus \bigcup_m^n B_m\right)\right) \cup \left(\left(B_n \setminus \bigcup_m^n B_m\right) \setminus A_n\right)\right) \\
&\leq \overline{\mu}\left((A_n \setminus B_n) \cup (B_n \setminus A_n)\right) \\
&= \overline{\mu}(A_n \Delta B_n) \\
&= 0.
\end{aligned}$$

We can therefore, without loss of generality, assume that B_n is pairwise disjoint. Now, extend the sequence B_n to an internal $*$-sequence, by the κ-extension principle (1.4.11). Note that $\overline{\mu}\left(\bigcup_n^m A_n\right) = \overline{\mu}\left(\bigcup_n^m B_n\right)$ for any $m \in \mathbb{N} \cup \{\mathbb{N}\}$, since

$$\begin{aligned}
\overline{\mu}\left(\bigcup_n^m (A_n \cup B_n)\right) &= \overline{\mu}\left(\left(\bigcup_n^m A_n\right) \cup \left(\bigcup_n^m (B_n \setminus A_n)\right)\right) \leq \overline{\mu}\left(\bigcup_n^m A_n\right) + \sum_n^m \overline{\mu}(B_n \setminus A_n) \\
&\leq \overline{\mu}\left(\bigcup_n^m A_n\right) + \sum_n^m \overline{\mu}(B_n \Delta A_n) \\
&= \overline{\mu}\left(\bigcup_n^m A_n\right) \leq \overline{\mu}\left(\bigcup_n^m (A_n \cup B_n)\right)
\end{aligned}$$

and similarly with the roles of A_n and B_n reversed. Now,

$$\sum_n^m \overline{\mu}(A_n) = \sum_n^m {}^\circ\mu(B_n) = {}^\circ \sum_n^m \mu(B_n) = {}^\circ\mu\left(\bigcup_n^m B_n\right) = \overline{\mu}\left(\bigcup_n^m A_n\right),$$

for each $m \in \mathbb{N}$. Furthermore, for each $m \in \mathbb{N}$ we have that

$$\mu\left(\bigcup_n^m B_n\right) - \frac{1}{m} \leq \overline{\mu}\left(\bigcup_n^m B_n\right) \leq \overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right) \leq \overline{\mu}\left(\bigcup_n^{\lambda} B_n\right),$$

for any $m \in \mathbb{N}$ and $\lambda \in \mathbb{N}_\infty$. It follows that the internal inequality

$$\mu\left(\bigcup_n^m B_n\right) - \frac{1}{m} \leq \overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right),$$

holds for all $m \in {}^*\mathbb{N}$ up to some $m = \kappa \in \mathbb{N}_\infty$, by Cauchy's principle (1.5.2). Thus $\mu\left(\bigcup_n^\lambda B_n\right) \simeq \overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right)$ for all $\lambda \in \mathbb{N}_\infty$ such that $\lambda \leq \kappa$, since we have that

$$\overline{\mu}\left(\bigcup_n^{\lambda} B_n\right) \simeq \mu\left(\bigcup_n^{\lambda} B_n\right) - \frac{1}{\lambda} \leq \overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right) \leq \overline{\mu}\left(\bigcup_n^{\lambda} B_n\right)$$

for all $\lambda \in \mathbb{N}_\infty$ such that $\lambda \leq \kappa$. Consequently we can conclude that

$$\sum_n^m \overline{\mu}(A_n) = {}^\circ\mu\left(\bigcup_n^m B_n\right) \to \overline{\mu}\left(\bigcup_n^{\mathbb{N}} B_n\right) = \overline{\mu}\left(\bigcup_n^{\mathbb{N}} A_n\right),$$

as $m \to \infty$ ($m \in \mathbb{N}$), by nonstandard convergence (1.6.8). □

In light of these results we make the following definition;

Definition 2.1.21 (Loeb measure)**.** The Loeb measure space $(\mu_L, \mathcal{A}_L, X)$ induced by $(\mu, \mathcal{A}, X)$ consists of the σ-algebra $\mathcal{A}_L = \Delta(\mu)$ and the measure $\mu_L = \overline{\mu}|_{\mathcal{A}_L}$.

All Loeb measures are complete in the usual sense.

Definition 2.1.22 (complete measure)**.** A measure $\mu : \mathcal{A} \to R$ on X is called complete if every subset of every measurable set with measure zero is measurable.

Proposition 2.1.23. *Loeb measures are complete.*

Proof. If $A \in \Delta(\mu)$ and $\mu_L(A) = 0$ and $B \subseteq A$ then clearly $B \in \Delta(\mu)$ and $\mu_L(B) = 0$. □

Proposition 2.1.24. *The μ-approximable sets are measurable.*

Proof. Let A be a μ-approximable set and B a μ_L-integrable set. It follows that

$$\mu_L(A \cap B) \leq \mu_L(B) < \infty.$$

Moreover, since A and B are μ-approximable it follows that there are sets $C, D \in \mathcal{A}$ such that $\mu_L(A\Delta C) = \mu_L(B\Delta D) = 0$. Now

$$\begin{aligned}
&\mu_L((A \cap B)\Delta(C \cap D)) \\
&\quad = \mu_L(((A \cap B) \setminus (C \cap D)) \cup ((C \cap D) \setminus (A \cap B))) \\
&\quad = \mu_L\left((A \cap B \cap (C \cap D)^{\mathsf{C}}) \cup (C \cap D \cap (A \cap B)^{\mathsf{C}})\right) \\
&\quad = \mu_L\left((A \cap B \cap (C^{\mathsf{C}} \cup D^{\mathsf{C}})) \cup (C \cap D \cap (A^{\mathsf{C}} \cup B^{\mathsf{C}}))\right) \\
&\quad = \mu_L\left((A \cap B \cap C^{\mathsf{C}}) \cup (A \cap B \cap D^{\mathsf{C}}) \cup (C \cap D \cap A^{\mathsf{C}}) \cup (C \cap D \cap B^{\mathsf{C}})\right) \\
&\quad \leq \mu_L\left((A \cap C^{\mathsf{C}}) \cup (B \cap D^{\mathsf{C}}) \cup (C \cap A^{\mathsf{C}}) \cup (D \cap B^{\mathsf{C}})\right) \\
&\quad = \mu_L((A\Delta C) \cup (B\Delta D)) \leq \mu_L(A\Delta C) + \mu_L(B\Delta C) = 0.
\end{aligned}$$

It follows that $A \cap B$ is μ_L-integrable. Thus A is μ_L-measurable. □

In the literature, a Loeb measure $(X, \mathcal{A}_L, \mu_L)$ is called hyperfinite if X is $*$-finite. We will instead opt for the convention that a Loeb measure is hyperfinite if the original internal algebra $\mathcal{A}$ is $*$-finite. Our definition clearly includes all the Loeb measures usually referred to as hyperfinite, since an algebra on a $*$-finite set must itself be $*$-finite. Hyperfinite Loeb measures have some properties that make them particularly desirable. In essence their desireable qualities derive from the fact that the $*$-finite algebra "behaves" as if it were a finite one. We will further elaborate on the value of hyperfiniteness in subsequent sections.

2.1.1 The Lebesgue Measure

As a first application of the theory of Loeb measures we shall construct the Lebesgue measure using the theory of Loeb measures. The obvious Loeb measure related to the Lebesgue measure is the Loeb measure $({}^*\nu)_L$, where ν is the Lebesgue measure. Roughly speaking we have the following equality

$$({}^*\nu)_L \circ {}^* = \nu.$$

But, nothing is really gained by doing this. The above equality is in fact true for any standard measure μ. And it is not hard to see why. For, the Loeb measure coincides with the inducing $*$-finitely additive measure on the elements of the algebra.

Instead we shall produce the Lebesgue measure from a hyperfinite Loeb measure. The idea is simply to take the infinitessimal intuition of integration and build upon it. Instead of defining our charge on the whole of ${}^*\mathbb{R}^n$, we will instead define it on an infinitely large cube containing all of $\mathbb{R}^n$. Let $N \in \mathbb{N}_\infty$ be the length of the sides of this cube, i.e. we will define our charge μ on $X = [-N, N)^n$. (The half-openness will make some technical details easier.) Next, we want to partition the space ${}^*[-N, N)^n$ into infinitesimal n-cubes. To do so we choose an infinitessimal length $\delta = \frac{1}{M} \simeq 0$, where $M \in \mathbb{N}_\infty$, as the length for the sides of our cubes. Our set of inifinitesimal cubes is then:

$$\mathcal{C} := \left\{ \prod_i^n [\eta(i)\delta, (\eta(i)+1)\delta) \middle| \eta \in ([-NM, MN) \cap {}^*\mathbb{Z})^n \right\}.$$

The set $\mathcal{C}$ has $(2MN)^n$ elements and, therefore, $*$-finite. The set algebra $\mathcal{A}$ it generates is therefore also $*$-finite. In particular, since the elements of $\mathcal{C}$ are pairwise disjoint every element of $\mathcal{A}$ can be expressed as a $*$-finite union of elements of $\mathcal{C}$.

Now, we define the internal charge $\mu : \mathcal{A} \to {}^*\mathbb{R}$ based on the idea that the total volume of each set $A \in \mathcal{A}$ is just the number of cubes in A times the volume of a cube, i.e.

$$\mu(A) = |A|\delta^n, \tag{2.3}$$

where $|A|$ denotes the $*$-number of cubes (elements of $\mathcal{C}$) in A. This is an instance of a *counting measure*, i.e. a measure which simply counts the elements (up to a scalar factor).

The induced Loeb measure has the following propety;

Proposition 2.1.25. *Each closed interval* $\prod_i^n {}^*[a_i, b_i]$, *where* $a_i, b_i \in \mathbb{R}$ *and* $n \in \mathbb{N}$, *is* μ_L-*measurable and*

$$\mu_L\left(\prod_i^n {}^*[a_i, b_i]\right) = \prod_i^n (b_i - a_i).$$

Proof. For any given $x \in {}^*\mathbb{R}$ let $\lfloor x \rfloor_\delta, \lceil x \rceil_\delta$ be the nearest lesser and greater $*$-integer multiple of δ respectively.

Now note that

$$\prod_i^n [\lceil a_i \rceil_\delta, \lfloor b_i \rfloor_\delta], \prod_i^n [\lfloor a_i \rfloor_\delta, \lceil b_i \rceil_\delta] \in \mathcal{C}$$

and

$$\prod_i^n [\lceil a_i \rceil_\delta, \lfloor b_i \rfloor_\delta] \subseteq \prod_i^n {}^*[a_i, b_i] \subseteq \prod_i^n [\lfloor a_i \rfloor_\delta, \lceil b_i \rceil_\delta].$$

Moreover,

$$\mu\left(\prod_i^n [\lceil a_i \rceil_\delta, \lfloor b_i \rfloor_\delta]\right) = \prod_i^n (\lfloor b_i \rfloor_\delta - \lceil a_i \rceil_\delta) \simeq \prod_i^n (b_i - a_i),$$

and similarly $\mu\left(\prod_i^n [\lfloor a_i \rfloor_\delta, \lceil b_i \rceil_\delta]\right) \simeq \prod_i^n (b_i - a_i)$. Thus,

$$\mu_L\left(\prod_i^n [\lceil a_i \rceil_\delta, \lfloor b_i \rfloor_\delta]\right) = \mu_L\left(\prod_i^n [\lfloor a_i \rfloor_\delta, \lceil b_i \rceil_\delta]\right) = \prod_i^n (b_i - a_i)$$

and in particular

$$\mu_L\left(\prod_i^n [\lceil a_i \rceil_\delta, \lfloor b_i \rfloor_\delta] \setminus \prod_i^n [\lfloor a_i \rfloor_\delta, \lceil b_i \rceil_\delta]\right) = 0.$$

Which proves that $\prod_i^n {}^*[a_i, b_i]$ is μ-approximable and thus μ_L-measurable, by definition. Moreover,

$$\mu_L\left(\prod_i^n {}^*[a_i, b_i]\right) = \prod_i^n (b_i - a_i).$$

□

Let $\nu : \mathcal{L} \to \mathbb{R}$ be the classical Lebesgue measure. The Lebesgue and Loeb measurable sets are then related in the following way;

Proposition 2.1.26. *If a set $C \subseteq \mathbb{R}^n$ is Lebesgue measurable then $st^{-1}[C] \cap X$ is Loeb measurable and*

$$\mu_L(st^{-1}[C] \cap X) = \nu(C).$$

Proof. Let C be a closed n-interval $\prod_i^n [a_i, b_i]$, $a_i, b_i \in \mathbb{R}^n$. Now,

$$\mathrm{st}^{-1}\left[\prod_i^n [a_i, b_i]\right] \cap X = \bigcap_m^{\mathbb{N}} \prod_i^n {}^*\left(\left[\left\lceil a_i - \frac{1}{m}\right\rceil_\delta, \left\lfloor b_i + \frac{1}{m}\right\rfloor_\delta\right]\right)$$

by Robinson's sequential lemma (1.6.7). Thus

$$\begin{aligned}
\mu_L\left(\mathrm{st}^{-1}\left[\prod_i^n [a_i, b_i]\right]\right) &= \mu_L\left(\bigcap_m^{\mathbb{N}} \prod_i^n {}^*\left[\left\lceil a_i - \frac{1}{m}\right\rceil_\delta, \left\lfloor b_i + \frac{1}{m}\right\rfloor_\delta\right]\right) \\
&= \lim_{m\to\infty} \mu_L\left(\prod_i^n {}^*\left[\left\lceil a_i - \frac{1}{m}\right\rceil_\delta, \left\lfloor b_i + \frac{1}{m}\right\rfloor_\delta\right]\right) \\
&= \lim_{m\to\infty} {}^\circ\prod_i^n \left(\left\lfloor b_i + \frac{1}{m}\right\rfloor_\delta - \left\lceil a_i - \frac{1}{m}\right\rceil_\delta\right) \\
&= \lim_{m\to\infty} \prod_i^n \left(b_i - a_i + \frac{2}{m}\right) \\
&= \prod_i^n (b_i - a_i) \\
&= \nu\left(\prod_i^n [a_i, b_i]\right)
\end{aligned}$$

by Proposition 2.1.25. It follows that $\mu_L \circ \mathrm{st}^{-1}$ and the Lebesgue measure agree on the Borel sets in general since st^{-1} commutes with set difference, countable union, and countable intersection. Moreover, it follows that they agree on the sets in the ν-completion of the Borel sets and thus they agree on all Lebesgue measurable sets.

□

2.2 Measurable and Internal Functions

In this section we will cover the basic nonstandard theory of measurable functions. We recall the following definitions.

Definition 2.2.1 (measurable)**.** A function $f : X \to Y$, where Y is a topological space, is called $\mathcal{A}$-measurable, where $\mathcal{A}$ is an algebra on X, if the preimage of any open set is an element of $\mathcal{A}$.

Definition 2.2.2 (simple)**.** An $\mathcal{A}$-measurable function is called simple if it only takes a finite number of values and $*$-simple if it only take a $*$-finite number of values.

Definition 2.2.3 (convergence in measure)**.** Let ν be a measure on a σ-algebra $\mathcal{A}$ on a set X, let $f_n : X \to \overline{\mathbb{R}}$ be a sequence of ν-measurable functions, and $f : X \to \overline{\mathbb{R}}$ a ν-measurable function. Then the sequence f_n is said to converge to f if for every $\delta \in \mathbb{R}_{>0}$ and any $\varepsilon \in \mathbb{R}_{>0}$ there exists a $n \in \mathbb{N}$ such that

$$\nu(\{x \in X \mid |f(x) - f_m(x)| \geq \delta) < \varepsilon,$$

for every $m \in \mathbb{N}_{\geq n}$.

As well as the following classical result.

Proposition 2.2.4. *If a sequence of functions converges to two functions in measure, the two functions must be equal almost everywhere. That is, if $f_n \xrightarrow{\mu} g$ and $f_n \xrightarrow{\mu} f$, then $f = g$ μ-almost everywhere.*

Proof. Let f, g and f_n be as in the statement of the proposition and let $\varepsilon \in \mathbb{R}_{>0}$. Then for each $n \in \mathbb{N}$ choose an $m_n \in \mathbb{N}$ such that

$$\mu\left(\left\{|f - f_{m_n}| \geq \frac{1}{2n}\right\}\right) < \frac{\varepsilon}{2^{n+1}} \quad \text{and} \quad \mu\left(\left\{|g - f_{m_n}| \geq \frac{1}{2n}\right\}\right) < \frac{\varepsilon}{2^{n+1}}.$$

It follows that f and g are equal μ-almost everywhere, since we have the inequality

$$\begin{aligned}
\mu(\{f \neq g\}) &= \mu\left(\bigcup_n^{\mathbb{N}} \left\{|f - g| \geq \frac{1}{n}\right\}\right) \leq \mu\left(\bigcup_n^{\mathbb{N}} \left\{|f - g| \geq \frac{1}{n}\right\}\right) \\
&\leq \sum_n^{\mathbb{N}} \mu\left(\left\{|f - g| \geq \frac{1}{n}\right\}\right) \leq \sum_n^{\mathbb{N}} \mu\left(\left\{|f - f_{m_n}| \geq \frac{1}{2n} \vee |g - f_{m_n}| \geq \frac{1}{2n}\right\}\right) \\
&\leq \sum_n^{\mathbb{N}} \left(\mu\left(\left\{|f - f_{m_n}| \geq \frac{1}{2n}\right\}\right) + \mu\left(\left\{|g - f_{m_n}| \geq \frac{1}{2n}\right\}\right)\right) \\
&< \sum_n^{\mathbb{N}} \frac{\varepsilon}{2^n} = 2\varepsilon,
\end{aligned}$$

and our choice of $\varepsilon \in \mathbb{R}_{>0}$ was arbitrary. □

We will now show the principal relations between the $\mathcal{A}$-measurable and $\mathcal{A}_L$-measurable functions. The results and their names come from *Foundations of Infinitesimal Stochastic Analysis* (Stroyan and Bayod 1986).) The statement of the uniform lifting lemma (2.2.7) has been modified, but is equivalent to the form presented by Stroyan and Bayod. The proof strategies are unchanged, however, we present the proofs in greater detail with minor deviations where we deemed it convenient.

We begin with the following definition.

Definition 2.2.5 (standard projection)**.** Let $f : X \to {}^*\mathbb{R}$ be an internal function, then the (standard) projection of f is the extended real function ${}^\circ f$.

The inverse of this notion in the context of an internal algebra $\mathcal{A}$ is;

Definition 2.2.6 (uniform lifting)**.** Let R be a topological space. Then an internal $\mathcal{A}$-measurable function $G : X \to R$ is called a *uniform lifting* of the function $g : X \to R$ if $^\circ G = g$.

The following proposition tells us precisely when a uniform lifting can be found. We cannot guarantee this in the general case where g is $\sigma(\mathcal{A})$-measurable, as this turns out to be insufficient.

Proposition 2.2.7 (uniform lifting lemma)**.** *Let R be a separable locally compact complete metric space and let S be a countable dense subset of R. Then a function $g : X \to R$ has an $\mathcal{A}$-uniform lifting $G : X \to {}^*R$ if and only if for each rational $r \in \mathbb{Q}$ and $s \in S$ the sets $g^{-1}[B(s,r)]$ are a countable union of elements in $\mathcal{A}$. Moreover, the lifting $G : X \to {}^*R$ can always be chosen to be $*$-simple.*

Proof. Suppose $G : X \to {}^*R$ is a $\mathcal{A}$-uniform lifting of $g : X \to R$. Note first that

$$x \in st^{-1}[B(s,r)] \Leftrightarrow {}^{\circ *}d(s,x) < r \Leftrightarrow \exists n \in \mathbb{N} \left({}^*d(s,x) < r - \frac{1}{n}\right)$$
$$\Leftrightarrow x \in \bigcup_n^{\mathbb{N}} {}^*B\left(s, r - \frac{1}{n}\right),$$

where $\mathrm{st} : {}^*\mathbb{R} \to \mathbb{R}$ is the standard part function, by Robinson's compactness criterion (1.6.11). It follows that for each $s \in S$ and $r \in \mathbb{Q}$ we have

$$g^{-1}[B(s,r)] = (\mathrm{st} \circ G)^{-1}[B(s,r)] = G^{-1}[\mathrm{st}^{-1}[B(s,r)]]$$
$$= G^{-1}\left[\bigcup_n^{\mathbb{N}} {}^*B\left(s, r - \frac{1}{n}\right)\right] = \bigcup_n^{\mathbb{N}} G^{-1}\left[B\left(s, r - \frac{1}{n}\right)\right].$$

Now since G is $\mathcal{A}$-measurable it follows that

$$\bigcup_n^{\mathbb{N}} G^{-1}\left[B\left(s, r - \frac{1}{n}\right)\right] \in \mathcal{A}_L.$$

Conversely, let $\bigcup_n^{\mathbb{N}} A_{s,r,n} = g^{-1}[B(s,r)]$ where $A_{s,r,n} \in \mathcal{A}$ for each $r \in \mathbb{Q}$, $s \in S$, and $n \in \mathbb{N}$. Let $C_{s,r,n} = \bigcup_m^n A_{s,r,m}$ and $E_{s,r,n}$ be the set of internal $\mathcal{A}$-measurable $*$-simple functions $f : X \to {}^*R$ such that

$$f[C_{s,r,n}] \subseteq {}^*B(s,r).$$

Note that each $E_{s,r,n}$ is nonempty, since it contains at least the function that is constantly equal to s. First suppose that the collection of sets $E_{s,r,n}$ has the finite intersection property. Then, since the collection is countable, it has a nonempty intersection, by the κ-intersection principle (1.4.10). Let $G : X \to {}^*R$ be in the intersection. By construction it is internal and $\mathcal{A}$-measurable. Now, for each $w \in X$ choose a countable basis $(B(s_{g(w),m}, r_{g(w),m}))_m^{\mathbb{N}}$ for the neighborhood system of $g(w)$ and let $N_{g(w)} \in \mathbb{N}$ be the index such that for all $n \geq N_{g(w)}$ we have $w \in C_{s_{g(w),m}, r_{g(w),m}, n}$. Now,

$$\mathrm{st}^{-1}[g(w)] = \bigcap_m^{\mathbb{N}} {}^*B\left(s_{g(w),m}, r_{g(w),m}\right) \supseteq \bigcap_m^{\mathbb{N}} G\left[C_{s_{g(w),m}, r_{g(w),m}, N_{g(w)}}\right]$$
$$\supseteq G\left[\bigcap_m^{\mathbb{N}} C_{s_{g(w),m}, r_{g(w),m}, N_{g(w)}}\right] \supseteq \{G(w)\}.$$

Thus $^\circ G = g$.

We now show that the collection of sets $E_{s,r,n}$ has the finite intersection property. For each $(x_0, x_1, x_2) = x \in S \times \mathbb{Q} \times \mathbb{N}$ let $C_x^0 = C_{x_0,x_1,x_2}$ and $C_x^1 = C_{x_0,x_1,x_2}^{\complement}$. Let H be a finite subset of $S \times \mathbb{Q} \times \mathbb{N}$. Now if an $\mathcal{A}$-measurable function G for each $c \in 2^H$ satisfies

$$G\left[\bigcap_h^H C_h^{c(h)}\right] \subseteq \bigcap_h^{\{h'|c(h')=0\}} {}^*B(h_0, h_1), \tag{2.4}$$

it follows that for each $(h_0, h_1, h_2) = h \in H$,

$$\begin{aligned} G\left[C_{h_0,h_1,h_2}\right] &= G[C_h] \\ &= G\left[\bigcup_c^{2^{H\setminus\{h\}}} \left(C_h^0 \cap \bigcap_{h'}^{H\setminus\{h\}} C_{h'}^{c(h')}\right)\right] = \bigcup_c^{2^{H\setminus\{h\}}} G\left[C_h^0 \cap \bigcap_{h'}^{H\setminus\{h\}} C_h'^{c(h')}\right] \\ &\subseteq \bigcup_c^{2^{H\setminus\{h\}}} \left(B(h_0, h_1) \cap \bigcap_{h'}^{\{h''|c(h'')=0\}} B(h_0', h_1')\right) \subseteq B(h_0, h_1). \end{aligned}$$

In particular, this means that $G \in E_{h_0,h_1,h_2}$ for each $h \in H$.

Finally, we just need to find an internal $\mathcal{A}$-measurable function satisfying (2.4). Firstly, we must show that if $\bigcap_h^H C_h^{c(h)} \neq \emptyset$ then $\bigcap_h^{c^{-1}[0]} {}^*B(h_0, h_1) \neq \emptyset$ so that functions consistent with (2.4) can exist. To see this note that

$$\emptyset \neq \bigcap_h^H C_h^{c(h)} \subseteq \bigcap_h^{\{h'|c(h')=0\}} C_h^0 \subseteq \bigcap_h^{\{h'|c(h')=0\}} g^{-1}[B(h_0, h_1)] = g^{-1}\left[\bigcap_h^{\{h'|c(h')=0\}} B(h_0, h_1)\right],$$

and so, by transfer, $\bigcap_h^{\{h'|c(h')=0\}} {}^*B(h_0, h_1) \neq \emptyset$. Now, observe that $\{\bigcap_h^H C_h^{c(h)} \mid c \in 2^H\}$ is a partition of X and that it is a subset of $\mathcal{A}$. Thus, the function G which given any $x \in \bigcap_h^H C_h^{c(h)}$ returns some fixed $G(x) \in \bigcap_h^{\{h'|c(h')=0\}} {}^*B(h_0, h_1)$ satisfies all the necessary criteria. It follows that the collection of the sets $E_{s,r,n}$ has the finite intersection property. □

This can be generalised to any first countable locally compact Hausdorff space;

Proposition 2.2.8 (uniform lifting lemma (locally compact))**.** *Let R be a first countable locally compact Hausdorff space and let S be a dense subset of R. Then a function $g : X \to R$ has a uniform lifting $G : X \to {}^*R$ if and only if for each $B \in \mathcal{B}_s$, $s \in S$ the sets $g^{-1}[B]$ are a countable union of elements in $\mathcal{A}$, where $\mathcal{B}_s$ is a countable neighbourhood basis for $\mathcal{T}_s$. Moreover, the lifting $G : X \to {}^*R$ can always be chosen to be $*$-simple.*

The major structure of the proof is unchanged. The first countability is just to ensure that the projection remains $\mathcal{A}_L$- (and $\sigma(\mathcal{A})$-) measurable. Separability is not needed since the saturation principle can handle any cardinality. The combined neighborhood systems for the dense subset then acts as a basis for the space as a whole space.

A more useful and forgiving notion of lifting is that of the μ-lifting;

Definition 2.2.9 (μ-lifting)**.** An internal $\mathcal{A}$-measurable function $G : X \to {}^*\mathbb{R}$ is called a *μ-lifting* of any function if $g : X \to \overline{\mathbb{R}}$ if ${}^\circ G = g$, μ_L-almost everywhere.

Proposition 2.2.10 (μ-lifting theorem)**.** *If $g : X \to \overline{\mathbb{R}}$ is $\mathcal{A}_L$-measurable and*

$$\mu_L\left(g^{-1}\left[\left(-\frac{1}{n}, \frac{1}{n}\right)^{\complement}\right]\right) < \infty,$$

for every $n \in \mathbb{N}$, then there exists an internal $*$-simple $\mathcal{A}$-measurable function $G : X \to {}^*\mathbb{R}$ such that ${}^\circ G = g$ μ_L-almost everywhere.

Proof. Let $D = \{d_n \in \overline{\mathbb{R}} \mid n \in \mathbb{N}\}$ be an indexed set which is dense in the range of g, that is $\operatorname{ran}(g) \subseteq \overline{D}$, and which includes ∞ and $-\infty$ as elements. Now define

$$M_n^m = \begin{cases} \left(-\frac{1}{m}, \frac{1}{m}\right)^{\complement} \cap \left(d_n - \frac{1}{m}, d_n + \frac{1}{m}\right), & \text{when } d_n \in \mathbb{R}, \\ \{\infty\}, & \text{when } d_n = \infty, \\ \{-\infty\}, & \text{when } d_n = -\infty, \end{cases}$$

and $F_n^m = g^{-1}[M_n^m]$ for each $n \in \mathbb{N}$. Finally, let $M_{-1}^m = \left(-\frac{1}{m}, \frac{1}{m}\right)$ and $F_{-1}^m = g^{-1}[M_{-1}^m]$.

Note that for any $m \in \mathbb{N}$ the set $M^m := \bigcup_n^{\mathbb{Z}_{\geq -1}} M_n^m$ covers the the range of g. For, suppose r is in the $\operatorname{ran}(g)$. If $r = \pm\infty$, then $r = d_n = \pm\infty$ for some $n \in \mathbb{N}$ and thus $r \in M_n^m$. If $r \in \left(-\frac{1}{m}, \frac{1}{m}\right) \cap \operatorname{ran}(g)$, then $r \in M_{-1}$. Finally, if $r \in \left(-\frac{1}{m}, \frac{1}{m}\right)^{\complement} \cap \operatorname{ran}(g)$, then there is a $d_n \in D$ for some $n \in \mathbb{N}$ such that $|r - d_n| < \frac{1}{m}$ and thus $r \in F_n^m$. It follows that $X = \bigcup_n^{\mathbb{Z}_{\geq -1}} F_n^m$ for any $m \in \mathbb{N}$, since for any $m \in \mathbb{N}$ we have that

$$X = g^{-1}[\operatorname{ran}(g)] \subseteq g^{-1}\left[\bigcup_n^{\mathbb{Z}_{\geq -1}} M_n^m\right] = \bigcup_n^{\mathbb{Z}_{\geq -1}} g^{-1}\left[M_n^m\right] = \bigcup_n^{\mathbb{Z}_{\geq -1}} F_n^m,$$

and $F_n^m \subseteq X$ for every $n \in \mathbb{Z}_{\geq -1}$.

Moreover, since g is $\mathcal{A}_L$-measurable and

$$\mu_L\left(\bigcup_n^{\mathbb{N}} F_n^m\right) \leq \mu_L\left(g^{-1}\left[\left(-\frac{1}{m}, \frac{1}{m}\right)^{\complement}\right]\right) < \infty,$$

for any given $m \in \mathbb{N}$, it follows that for any given $m \in \mathbb{N}$ there exists a $h \in \mathbb{N}$ large enough that $\mu_L\left(\bigcup_n^{\mathbb{N}_{\geq h}} F_n^m\right) < \frac{1}{m}$. Furthermore, we can choose a set $A_n^m \in \mathcal{A}$ for each $n \in h$ such that $\mu_L\left(F_n^m \setminus A_n^m\right) \leq \frac{1}{mh}$, by the flanking characterisation of $\Delta(\mu)$ (2.1.13), since each F_n^m is μ_L-integrable and thus μ-approximable. In particular it follows that $\sum_n^h \mu_L\left(F_n^m \setminus A_n^m\right) < \frac{1}{m}$ for any $m \in \mathbb{N}$.

Now we can partition the set $A^m = \bigcup_n^h A_n^m$ by intersections in the following way

$$P^m = \left\{\bigcap_n^h A_{n,c(n)}^m \neq \emptyset \middle| c \in 2^h\right\},$$

where $A_{n,0}^m = A_n^m$ and $A_{n,1}^m = A^m \setminus A_n^m$. Since each $B \in P^m$ is a subset of some F_n^m, $n \in h$, we can choose $\eta : P^m \to \{n \mid n \in h\}$ such that if $\eta(B) = n$ then $B \subseteq F_n^m$ for each $B \in P^m$. In short, $B \subseteq F_{\eta(B)}^m$. This lets us define an internal $*$-simple $\mathcal{A}$-measurable function $G_m : X \to \{d_n \mid n \in h\}$,

$$G_m(x) := \begin{cases} d_{\eta(B)}, & \text{if } x \in B \in P^m, \\ 0, & \text{if } x \notin A^{\complement}. \end{cases}$$

For the remainder of this proof we adopt the convention that $|\infty - \infty| = 0$. It follows that

$$|G_m(x) - g(x)| < \frac{1}{m}, \quad \forall x \in A^m \cup F_{-1}^m. \tag{2.5}$$

For, suppose $x \in A^m$. Then $x \in B$ for some unique $B \in P^m$, since P^m partitions A^m. Consequently, $x \in B \subseteq F_{\eta(B)}^m$ as noted earlier and thus

$$\begin{aligned} x \in F_{\eta(B)}^m &= g^{-1}\left[\left(-\frac{1}{m}, \frac{1}{m}\right)^{\complement} \cap \left(d_{\eta(B)} - \frac{1}{m}, d_{\eta(B)} + \frac{1}{m}\right)\right] \\ &\subseteq g^{-1}\left[\left(d_{\eta(B)} - \frac{1}{m}, d_{\eta(B)} + \frac{1}{m}\right)\right] \\ &= g^{-1}\left[\left(G_m(x) - \frac{1}{m}, G_m(x) + \frac{1}{m}\right)\right], \end{aligned}$$

when $d_{\eta(B)} \in \mathbb{R}$ and

$$x \in F^m_{\eta(B)} = g^{-1}[\{d_{\eta(B)}\}] = g^{-1}[\{G_m(x)\}],$$

when $d_{\eta(B)} \in \{-\infty, \infty\}$.

Suppose now that $x \in F^m_{-1} = g^{-1}\left[\left(-\frac{1}{m}, \frac{1}{m}\right)\right]$. Then $x \notin g^{-1}\left[\left(-\frac{1}{m}, \frac{1}{m}\right)^{\complement}\right] \supseteq F^m_n$ for any $n \in \mathbb{N}$. In particular, this means that $x \notin \bigcup P^m = A^m$, since every $B \in P^m$ is in some F^m_n, $n \in h$. That is x is in the complement of A^m and thus $G_m(x) = 0$. It follows that $|G_m(x) - g(x)| = |g(x)| \in \left(-\frac{1}{m}, \frac{1}{m}\right)$. Inequality (2.5) follows.

It in turn implies that

$$\mu_L\left(\left\{x \in X \middle| |G_m(x) - g(x)| \geq \frac{1}{m}\right\}\right) \leq \frac{2}{m}.$$

In fact, we have

$$\begin{aligned}
&\mu_L\left(\left\{x \in X \middle| |G_m(x) - g(x)| \geq \frac{1}{m}\right\}\right) = \mu_L\left(\left\{x \in X \middle| |G_m(x) - g(x)| < \frac{1}{m}\right\}^{\complement}\right) \\
&\leq \mu_L\left(\left(A^m \cup F^m_{-1}\right)^{\complement}\right) = \mu_L\left(\left(\bigcup_n^h A^m_n \cup F^m_{-1}\right)^{\complement}\right) \\
&= \mu_L\left(\bigcap_n^h (A^m_n)^{\complement} \cap (F^m_{-1})^{\complement}\right) = \mu_L\left(\left(F^m_{-1} \cup \bigcup_n^{\mathbb{N}} F^m_n\right) \cap \bigcap_n^h (A^m_n)^{\complement} \cap (F^m_{-1})^{\complement}\right) \\
&= \mu_L\left(\left(\bigcup_n^{\mathbb{N}} F^m_n\right) \cap \bigcap_n^h (A^m_n)^{\complement} \cap (F^m_{-1})^{\complement}\right) = \mu_L\left(\left(\bigcup_n^{\mathbb{N}} F^m_n\right) \cap \bigcap_n^h (A^m_n)^{\complement}\right) \\
&= \mu_L\left(\bigcup_n^{\mathbb{N}}\left(F^m_n \cap \bigcap_k^h (A^m_k)^{\complement}\right)\right) \leq \mu_L\left(\left(\bigcup_n^{\mathbb{N}\setminus h} F^m_n\right) \cup \left(\bigcup_n^h F^m_n \cap (A^m_n)^{\complement}\right)\right) \\
&= \mu_L\left(\left(\bigcup_n^{\mathbb{N}_{\geq h}} F^m_n\right) \cup \left(\bigcup_n^h (F^m_n \setminus A^m_n)\right)\right) \leq \mu_L\left(\bigcup_n^{\mathbb{N}_{\geq h}} F^m_n\right) + \sum_n^h \mu_L\left(F^m_n \setminus A^m_n\right) \\
&\leq \frac{2}{m}.
\end{aligned}$$

We can now observe that the sequence ${}^\circ G_m$, $m \in \mathbb{N}$ converges to f in measure, since

$$\begin{aligned}
&\mu_L\left(\left\{x \in X \middle| |{}^\circ G_m(x) - g(x)| \geq \frac{2}{m}\right\}\right) \\
&\leq \mu_L\left(\left\{x \in X \middle| |{}^\circ G_m(x) - G_m(x)| \geq \frac{1}{m} \vee |G_m(x) - g(x)| \geq \frac{1}{m}\right\}\right) \\
&\leq \mu_L\left(\left\{x \in X \middle| |G_m(x) - f(x)| \geq \frac{1}{m}\right\}\right) \\
&\leq \frac{2}{m}.
\end{aligned}$$

Now extend G_n to an internal $*$-sequence of $*$-simple functions, by the κ-extension principle (1.4.11). Observe that since the inequality

$$\begin{aligned}
&\mu\left(\left\{x \in X \middle| |G_m(x) - G_n(x)| \geq \frac{2}{m}\right\}\right) < \mu_L\left(\left\{x \in X \middle| |G_m(x) - G_n(x)| \geq \frac{2}{m}\right\}\right) + \frac{1}{m} \\
&\leq \mu_L\left(\left\{x \in X \middle| |G_m(x) - g(x)| \geq \frac{1}{m} \vee |G_n(x) - g(x)| \geq \frac{1}{m}\right\}\right) + \frac{1}{m} \\
&\leq \mu_L\left(\left\{x \in X \middle| |G_m(x) - g(x)| \geq \frac{1}{m}\right\}\right) + \mu_L\left(\left\{x \in X \middle| |G_n(x) - g(x)| \geq \frac{1}{m}\right\}\right) + \frac{1}{m} \\
&\leq \frac{2}{m} + \frac{2}{n} + \frac{1}{m} \leq \frac{5}{m},
\end{aligned}$$

holds for any $m, n \in \mathbb{N}$ such that $m < n$, it follows that the internal formula

$$\mu\left(\left\{x \in X \middle| |G_m(x) - G_n(x)| \geq \frac{2}{m}\right\}\right) \leq 5m, \quad \forall m \in {}^*\mathbb{N}_{<n},$$

is true for all $n \in \mathbb{N}$. Thus it holds up to some $n = \kappa \in \mathbb{N}_\infty$, by Cauchy's principle (1.5.2).

Finally, we observe that ${}^\circ G_n$, $n \in \mathbb{N}$ also converges to ${}^\circ G_\kappa$ in measure, since

$$\begin{aligned}
&\mu_L\left(\left\{x \in X \middle| |{}^\circ G_\kappa(x) - {}^\circ G_m(x)| \geq \frac{2}{m}\right\}\right) \\
&\quad \leq \mu_L\left(\left\{x \in X \middle| |{}^\circ G_\kappa(x) - G_\kappa(x)| + |G_\kappa(x) - G_m(x)| + |G_m(x) - {}^\circ G_m(x)| \geq \frac{2}{m}\right\}\right) \\
&\quad \leq \mu_L\left(\left\{x \in X \middle| |G_\kappa(x) - G_m(x)| \geq \frac{1}{m}\right\}\right) \\
&\quad \leq \frac{5}{m}.
\end{aligned}$$

It follows that ${}^\circ G_\kappa$ is μ_L-almost everywhere equal to g, by Proposition 2.2.4. Since G_κ is $*$-simple and $\mathcal{A}$-measurable we can define a $*$-simple $H_\kappa : X \to {}^*\mathbb{R}$ with the same standard part as G_κ. For instance, we can use the following scheme

$$H_\kappa(x) = \begin{cases} G_\kappa(x), & \text{when } G_\kappa(x) \notin \{-\infty, \infty\} \\ \kappa, & \text{when } G_\kappa(x) = \infty, \\ -\kappa, & \text{when } G_\kappa(x) = -\infty. \end{cases}$$

Clearly ${}^\circ G_\kappa = {}^\circ H_\kappa$ and thus ${}^\circ H_\kappa : X \to$ is equal to g μ_L-almost everywhere. □

2.3 Theory of Integration

The integration theory we will cover here will show that, in a sense, the $*$-simple functions are sufficient for a full integration theory. [1] If $f : X \to {}^*\mathbb{R}$ is a $*$-real $*$-simple $\mathcal{A}$-measurable function it can be expressed as follows;

$$f = \sum_{v}^{\mathcal{R}(f)} \chi_{f^{-1}[\{v\}]},$$

where $\mathcal{R}(f)$ is the range of f. The integral for $*$-simple functions is defined, in the usual way;

Definition 2.3.1. Let $f : X \to {}^*\mathbb{R}$ be an μ-measurable $*$-simple function then the integral is defined

$$\int f d\mu = \sum_{v}^{\mathcal{R}(f)} v\mu(f^{-1}[\{v\}]).$$

The $*$-simple $*$-real valued functions form a $*$-linear space with respect to which the integral is $*$-linear as well. Moreover the integral is monotone, i.e. if f, g are $*$-simple μ-measurable functions such that $f < g$ then $\int f d\mu < \int g d\mu$.

We will need the following proposition of $\mathcal{A}$-measurable functions;

Proposition 2.3.2. *If $f : X \to Y$ is an $\mathcal{A}$-measurable function, where $\mathcal{A}$ is the algebra generated by the topology on X, and $g : Y \to Z$ is a continuous function, then $g \circ f$ is $\mathcal{A}$-measurable.*

Proof. If C is an open set in Z, then ü

$$(g \circ f)^{-1}[C] = f^{-1}\left[g^{-1}[C]\right].$$

Now, since g is continuous, it follows that $g^{-1}[C]$ is open and thus $f^{-1}\left[g^{-1}[C]\right]$ is in $\mathcal{A}$. □

In particular, if f is $\mathcal{A}$-measurable it follows that $|f|$ is $\mathcal{A}$-measurable, since the absolute value is continuous. Notably, it applies to the $*$-simple $*$-real valued functions, i.e. if $f : X \to {}^*\mathbb{R}$ is $*$-simple then so is $|f|$.

In practice, when Loeb measures are used one almost always works with hyperfinite Loeb spaces, i.e. $\mathcal{A}$ is $*$-finite. In these spaces all $\mathcal{A}$-measurable $*$-real valued functions are $*$-simple. Moreover, as we saw in 2.2 Measurable and Internal Functions we can always choose the lifting to be $*$-simple and so the $*$-simple $\mathcal{A}$-measurable functions are in a sense sufficient. Therefore, for the rest of this section the reader may assume that we are working with a hyperfinite Loeb measure. In particular, from this point on μ-measurable and $\mathcal{A}$-measurable functions are assumed to be $*$-simple.

In the remainder of this section we cover the theory of S-integrability. Which establishes a correspondence between the Loeb integral (integral with respect to μ_L) and the integral on the inducing measure (integral with respect to μ). The section is primarily based on the treatment of this subject in *Foundations of Infinitesimal Stochastic Analysis* (Stroyan and Bayod 1986). Although we extend it to the context of a hyperfinite unbounded Loeb measure.

Definition 2.3.3 (S-integrable)**.** An internal μ-measurable function $f : X \to {}^*\mathbb{R}$ is called *S-integrable* if it has the following properties:

$$^\circ\int |f| d\mu < \infty, \qquad \text{(finite integral)}$$

$$\text{if } A \in \mathcal{A} \text{ and } \mu(A) \simeq 0 \text{ then } \int_A |f| \simeq 0, \qquad \text{(S-absolute continuity)}$$

$$\text{if } A \in \mathcal{A} \text{ and } f|_A \simeq 0 \text{ then } \int_A |f| \simeq 0. \qquad \text{(S-monotone continuity)}$$

[1]The treatment can be extended to a more general class of functions but this is beyond the scope of this work. The issue one must get around is that $*$-real $\mathcal{A}$-measurable functions are not, in general, closed under addition when $\mathcal{A}$ is an algebra.

Here $f|_A \simeq 0$ means that the restriction of f to the set A only takes on infinitesimal values.

In case the measure μ (or equivalently μ_L) is bounded the third criterion is redundant. Since, if $|f| \simeq 0$ on $A \in \mathcal{A}$ then $\int_A |f| d\mu \le \int_A \varepsilon d\mu = \varepsilon \mu(A)$ for each $\varepsilon \in \mathbb{R}_{>0}$. Which shows that $\int_A |f| d\mu \simeq 0$ since $\mu(A)$ is finite.

Proposition 2.3.4. *A function $f : X \to {}^*\mathbb{R}$ is S-integrable if and only if it has the following properties:*

$$\begin{aligned} &{}^\circ\int |f| d\mu \ < \infty, && \\ &\int_{|f|^{-1}[0,\frac{1}{N}]} |f| d\mu \simeq 0, && \forall N \in \mathbb{N}_\infty, \\ &\int_{|f|^{-1}[N,{}^*\infty]} |f| d\mu \simeq 0, && \forall N \in \mathbb{N}_\infty. \end{aligned}$$

Proof. Suppose f is S-integrable then the three properties follow, by definition.

Conversely, suppose f satisfies the properities outlined in the statement. Suppose $A \in \mathcal{A}$ is such that $\mu(A) \simeq 0$, then choose $K \in \mathbb{N}_\infty$ such that $K\mu(A) \simeq 0$. Now,

$$\begin{aligned} \int_A |f| d\mu &= \int_{A \setminus |f|^{-1}[K,{}^*\infty]} |f| d\mu + \int_{A \cap |f|^{-1}[K,{}^*\infty]} |f| d\mu \\ &\le \int_A K d\mu + \int_{|f|^{-1}[K,{}^*\infty]} |f| d\mu \\ &= K\mu(A) + \int_{|f|^{-1}[K,{}^*\infty]} |f| d\mu \\ &\simeq 0. \end{aligned}$$

Suppose now that $A \in \mathcal{A}$ such that $f|_A \simeq 0$. Note that $f|_A < \frac{1}{N}$ for all $N \in \mathbb{N}$, and therefore this is true up to some $N \in \mathbb{N}_\infty$, by Cauchy's principle (1.5.2). Thus,

$$\int_A |f| d\mu \le \int_{|f|^{-1}[0,\frac{1}{N}]} |f| d\mu \simeq 0.$$

□

Naturally, the third condition is again redundant in the case of a bounded measure. The following proposition gives us a fundamental class of S-integrable functions.

Proposition 2.3.5 (finite function S-integrability)**.** *Every internal $\mathcal{A}$-measurable function which has μ-finite support and is absolutely bounded μ-a.e. is S-integrable.*

Proof. Let f be a function as in the proposition. Let M be a finite upper bound of the range of $|f|$ and let $N \in {}^*\mathbb{R}_{\not\simeq\infty}$ be the measure of the support of f. Then cleary,

$${}^\circ\int |f| d\mu \le MN < \infty.$$

Furthermore for each $A \in \mathcal{A}$ such that $\mu(A) \simeq 0$ we have

$$\int_A |f| d\mu \le M\mu(A) \simeq 0.$$

And finally for each $A \in \mathcal{A}$ such that $f|_A \simeq 0$ we have

$$\int_A |f| d\mu \leq \varepsilon N,$$

for all $\varepsilon \in \mathbb{R}$ and so $\int_A |f| d\mu \simeq 0$. □

It turns out that S-integrability has the useful property of being absolutely downward closed;

Proposition 2.3.6 (dominated S-integrability criterion)**.** *If $f, g : X \to {}^*\mathbb{R}$ are $\mathcal{A}$-measurable, g is S-integrable, and $|f| \leq |g|$, then f is S-integrable.*

Proof. If $f, g : X \to {}^*\mathbb{R}$ are $\mathcal{A}$-measurable, g is S-integrable, and $|f| \leq |g|$. Obviously, $\int |f| \leq \int |g| < \infty$ and thus $\int |f| < \infty$. Consequently, f meets the first criterion of being S-integrable.

Now if $A \in \mathcal{A}$ and $\mu(A) \simeq 0$, we have $\int_A |f| \leq \int_A |g| \simeq 0$. Finally, let $\varepsilon \in \mathbb{R}_{>0}$, $A \in \mathcal{A}$ such that $f[A] \simeq 0$, then choose $n \in \mathbb{N}$ such that

$$\int_{|g|^{-1}[0,\frac{1}{n}]} |g| < \varepsilon.$$

Note that

$$m \int_{A \setminus |g|^{-1}[0,\frac{1}{n}]} |f| d\mu \leq \int_{A \setminus |g|^{-1}[0,\frac{1}{n}]} |g| < \infty$$

for every $m \in \mathbb{N}$, and thus $\int_{A \setminus |g|^{-1}[0,\frac{1}{n}]} |f| \simeq 0$. It follows that

$$\begin{aligned} \int_A |f| = \int_{A \cap |g|^{-1}[0,\frac{1}{n}]} |f| + \int_{A \setminus |g|^{-1}[0,\frac{1}{n}]} |f| \leq \int_{|g|^{-1}[0,\frac{1}{n}]} |g| + \int_{A \setminus |g|^{-1}[0,\frac{1}{n}]} |f| \\ < \varepsilon. \end{aligned}$$

And thus $\int_A |f| \simeq 0$. □

Proposition 2.3.7. *The S-integrable functions are closed under linear combination.*

Proof. Trivial. □

For the next proposition we will need the definition;

Definition 2.3.8. A sequence of μ-measurable functions f_n, $n \in \mathbb{N}$, is called S-Cauchy in the mean if for each $\varepsilon \in \mathbb{R}_{>0}$ there exists an $m \in \mathbb{N}$ such that for all $h, k \geq m$, $k, k \in \mathbb{N}$,

$$\int |f_h - f_k| d\mu < \varepsilon.$$

Proposition 2.3.9. *Every sequence of S-integrable which is S-Cauchy in the mean has an S-integrable S-limit in the mean.*

Proof. Let f_n be a sequence of S-integrable functions which is S-Cauchy in the mean. Extend the sequence to an internal $*$-sequence, by the κ-extension principle (1.4.11). Now each set

$$E_n := \left\{ m \in {}^*\mathbb{N} \middle| \forall k, j \in \mathbb{N}_{\geq m} \left(\int |f_k - f_j| < \frac{1}{n} \text{ and } f_k, f_j \text{ are } \mathcal{A}\text{-measurable} \right) \right\},$$

where $n \in \mathbb{N}$, is internal and contains a terminal segment of $\mathbb{N}$, since the sequence is Cauchy on the natural indices. It follows that each E_n contains an initial segment of $\mathbb{N}_\infty$, by Cauchy's principle (1.5.2). It follows that the collection of sets E_n has the finite intersection property, and therefore their intersection ($\bigcap_n^{\mathbb{N}} E_n$) is nonempty, by the κ-intersection principle (1.4.10). Let $\kappa \in \bigcap_n^{\mathbb{N}} E_n$.

Then for each $n \in \mathbb{N}$ there is an $m \in \mathbb{N}$ such that for all $h \in \mathbb{N}_{\geq m}$ we have $\int |f_\kappa - f_h| d\mu < \frac{1}{n}$. In particular $\lim_{m\to\infty} {}^\circ \int |f_\kappa - f_m| = 0$. If $\kappa \in \mathbb{N}$ then we are done.

In the case $\kappa \in \mathbb{N}_\infty$ it still remains to be shown that f_κ is S-integrable. Firstly,

$$\int |f_\kappa| d\mu \leq \int |f_\kappa - f_m| d\mu + \int |f_m| d\mu \leq \frac{1}{n} + \int |f_m| d\mu < \infty,$$

for large enough $m \in \mathbb{N}$. Secondly, for each $A \in \mathcal{A}$ such that $\mu(A) \simeq 0$ we have for each $n \in \mathbb{N}$,

$$\int_A |f_\kappa| d\mu \leq \int_A |f_\kappa - f_m| d\mu + \int_A |f_m| d\mu \leq \frac{1}{n} + \int_A |f_m| d\mu < \frac{1}{n},$$

and thus $\int_A |f_k| d\mu \simeq 0$. Finally, let $A \in \mathcal{A}$ such that $f|_A \simeq 0$. Note that in this case

$$m \int_{A \setminus f_m^{-1}[0,\frac{1}{h}]} |f_\kappa| d\mu \leq \int_{A \setminus f_m^{-1}[0,\frac{1}{h}]} |f_m| d\mu < \infty,$$

for each $m, h \in \mathbb{N}$, and thus $\int_{A \setminus f_m^{-1}[0,\frac{1}{h}} |f_\kappa| d\mu \simeq 0$. It follows that for each $n \in \mathbb{N}$

$$\begin{aligned}
\int_A |f_\kappa| d\mu &= \int_{A \setminus f^{-1}[0,\frac{1}{h}]} |f_\kappa| d\mu + \int_{A \cap f^{-1}[0,\frac{1}{h}]} |f_\kappa| d\mu \\
&\simeq \int_{A \cap f^{-1}[0,\frac{1}{h}]} |f_\kappa| d\mu \\
&\leq \int_{A \cap f^{-1}[0,\frac{1}{h}]} |f_\kappa - f_m| d\mu + \int_{A \cap f^{-1}[0,\frac{1}{h}]} |f_m| d\mu \\
&\leq \frac{1}{n} + \frac{1}{n},
\end{aligned}$$

for large enough $h, m \in \mathbb{N}$. And so $\int_A |f_\kappa| d\mu \simeq 0$. □

It turns out that the $\mathcal{A}$-measurable bounded functions with μ-finite support are S-dense in the set of S-integrable functions.

Proposition 2.3.10. *Every S-integrable function is an S-limit in the mean of a sequence of S-integrable functions which are bounded and have μ-finite support.*

Proof. Let f be an S-integrable function. Consider the internal $*$-sequence of functions defined by

$$f_n := \chi_{|f|^{-1}[\frac{1}{n},n]} f,$$

for each $n \in {}^*\mathbb{N}$. Then $\mu\left(|f|^{-1}\left[\frac{1}{n}, n\right]\right) < \infty$, since

$$\mu\left(|f|^{-1}\left[\frac{1}{n}, n\right]\right) \leq \int_{|f|^{-1}[\frac{1}{n},n]} n|f| d\mu \leq n \int |f| d\mu < \infty.$$

Thus for each $n \in \mathbb{N}$ the function f_n has μ-finite support. Furthermore since $f_n \leq f$ for each $n \in {}^*\mathbb{N}$, every f_n is S-integrable, by the dominated S-integrability criterion (2.3.6). Now

$$\begin{aligned}
\int |f - f_N| d\mu &= \int_{|f|^{-1}[0,\frac{1}{N}]} |f - f_N| d\mu + \int_{|f|^{-1}[\frac{1}{N},N]} |f - f_N| d\mu + \int_{|f|^{-1}[N,\infty]} |f - f_N| d\mu \\
&= \int_{|f|^{-1}[0,\frac{1}{N}]} |f| d\mu + \int_{|f|^{-1}[\frac{1}{N},N]} |f - f_N| d\mu + \int_{|f|^{-1}[N,\infty]} |f| d\mu \\
&= \int_{|f|^{-1}[0,\frac{1}{N}]} |f| d\mu + \int_{|f|^{-1}[N,\infty]} |f| d\mu \\
&\simeq 0,
\end{aligned}$$

for all $N \in \mathbb{N}_\infty$. And so f is the S-limit in the mean of f_n $(n \in \mathbb{N})$, by nonstandard convergence (1.6.8). □

The essential feature that makes S-integrable functions interesting is the following commutation relation relating the μ-integral and the Loeb integral;

Proposition 2.3.11. *If a function f is S-integrable then*

$$^\circ\int_A f d\mu = \int_A {}^\circ f d\mu_L < \infty,$$

for all $A \in \mathcal{A}$.

Proof. We begin, by observing that the proposition holds for any characteristic function of any member of $\mathcal{A}$ of finite μ-measure. By linearity this property extends to any linear combination of such characteristic functions.

We now prove the proposition for S-integrable functions which are absolutely bounded and have support which have finite μ-measure. Suppose f is such a function and let $M, N \in \mathbb{Z}$, $K \in \mathbb{N}$ such that M is the least integer upper bound of f, N is the greatest integer lower bound of f, and K is larger than the measure of the support of f. Define

$$f_0 := \sum_m^{[K,N)\cap^*\mathbb{Z}} m\chi_{f^{-1}[m,m+1)}$$
$$f_{n+1} := f_n + \frac{1}{2^{n+1}}\chi_{|f-f_n|^{-1}[\frac{1}{2^{n+1}},1]}.$$

for each $n \in \mathbb{N}^*$. Now note that ${}^\circ f_n \nearrow {}^\circ f$ everywhere, for $n \in \mathbb{N}$ and therefore

$$\lim_{n\to\infty}\int {}^\circ f_n d\mu_L = \int \lim_{n\to\infty} {}^\circ f_n d\mu_L = \int {}^\circ f d\mu_L,$$

by the monotone convergence theorem. Note that,

$$^\circ\int f_n d\mu = \int {}^\circ f_n d\mu_L,$$

by the first part of this proof. It follows that ${}^\circ\int f_n d\mu = \int {}^\circ f_n d\mu_L \to \int {}^\circ f d\mu_L$ as $n \to \infty$. But, we also have

$$|\int f d\mu - \int f_n d\mu| \leq \int |f - f_n| d\mu \leq \frac{K}{2^n} \simeq 0,$$

for all $n \in \mathbb{N}_\infty$. It follows that ${}^\circ\int f_n d\mu \to {}^\circ\int f d\mu$ as $n \to \infty$, by nonstandard convergence (1.6.8).

Finally, let f be a generic S-integrable function, i.e. possbily absolutely unbounded and with support possibly of infinite μ-measure. Now for any given $\varepsilon \in \mathbb{R}_{>0}$ for large enough $n \in \mathbb{N}$ we have $\int_{|f|^{-1}[0,\frac{1}{n})} |f| d\mu, \int_{|f|^{-1}(n,\infty]} |f| d\mu < \varepsilon$, and

$$\int_{|f|^{-1}(\frac{1}{n},n)} {}^\circ f d\mu_L < \int {}^\circ f d\mu_L + \varepsilon.$$

(The last one follows from the dominated convergence theorem.) Now on the one hand we have

$$\begin{aligned}
{}^\circ\int f d\mu &= {}^\circ\int_{|f|^{-1}[0,\frac{1}{n}]} f d\mu + {}^\circ\int_{|f|^{-1}(\frac{1}{n},n)} f d\mu + {}^\circ\int_{|f|^{-1}[n,\infty]} f d\mu \\
&\le \varepsilon + {}^\circ\int_{|f|^{-1}(\frac{1}{n},n)} f d\mu + \varepsilon \\
&\le 2\varepsilon + \int_{|f|^{-1}(\frac{1}{n},n)} {}^\circ f d\mu_L \\
&\le 2\varepsilon + \varepsilon + \int {}^\circ f d\mu_L.
\end{aligned}$$

Thus, ${}^\circ\int f d\mu \le \int {}^\circ f d\mu_L$. On the other hand it follows, by the first part of the argument, that

$$\begin{aligned}
\int {}^\circ f d\mu_L &= \lim_{n\to\infty} \int {}^\circ(\chi_{|f|^{-1}(\frac{1}{n},n)} f) d\mu_L \\
&= \lim_{n\to\infty} {}^\circ\int \chi_{|f|^{-1}(\frac{1}{n},n)} f d\mu \\
&\le {}^\circ\int f d\mu.
\end{aligned}$$

It follows that ${}^\circ\int f d\mu = \int {}^\circ f d\mu_L$. □

Finally, to finish the subject of this section we prove that each μ_L-integrable function has an S-integrable μ-lifting. This is essentially the converse of the previous theorem.

Proposition 2.3.12. *If $f : X \to \mathbb{R}$ is μ_L-integrable then it has an S-integrable μ-lifting $F : X \to {}^*\mathbb{R}$ such that ${}^\circ\int_A F d\mu = \int_A f d\mu_L$ for all $A \in \mathcal{A}$.*

Proof. The existence of a μ-lifting F follows from the μ-lifting theorem (2.2.10). Consider the $*$-sequence

$$F_n := \chi_{|F|^{-1}(\frac{1}{n},n)} F,$$

for $n \in {}^*\mathbb{N}_{\ge 2}$.

Clearly each F_n, $n \in \mathbb{N}_{\ge 2}$, is bounded. Moreover, the set $|F|^{-1}\left[{}^*\left(\frac{1}{n},n\right)\right]$ for each $n \in \mathbb{N}_{\ge 2}$ must have finite μ-measure. For, suppose $x \in X$ is such that ${}^\circ F(x) = f(x)$ and $x \in |F|^{-1}\left[{}^*\left(\frac{1}{n},n\right)\right]$, for some $n \in \mathbb{N}_{\ge 2}$. It follows that,

$$|f(x)| = |{}^\circ F(x)| \in \left[\frac{1}{n}, n\right]$$

and thus $x \in |f|^{-1}\left[\left(\frac{1}{2n}, 2n\right)\right]$. In particular,

$$\begin{aligned}
\mu_L\left(|F|^{-1}\left[\left(\frac{1}{n},n\right)\right]\right) &\le \mu_L\left(|f|^{-1}\left[\left(\frac{1}{2n},2n\right)\right] \cup \{{}^\circ F \ne f\}\right) \\
&\le \mu_L\left(|f|^{-1}\left[\left(\frac{1}{2n},2n\right)\right]\right) \\
&< \infty.
\end{aligned}$$

Consequently, F_n for each $n \in \mathbb{N}_{\ge 2}$ is S-integrable, by finite function S-integrability (2.3.5).

Next note that $\mu_L(\{|f| = \infty\}) = 0$, since f is μ_L-integrable.

Moreover, note that if $x \in X$ is such that ${}^\circ F = f(x)$ and $|f(x)| < \infty$, then ${}^\circ F_n(x) \to f(x)$. For, if $f(x) = 0$, then ${}^\circ F_n(x) = f(x)$ for all $n \in \mathbb{N}_{\ge 2}$. On the other hand, if $f(x) \ne 0$ then $|f(x)| \in \left(\frac{1}{n}, n\right)$ for some $n \in \mathbb{N}_{\ge 2}$. Consequently, $f(x) = {}^\circ F_m(x)$ for all $m \in \mathbb{N}_{2n}$.

It follows that ${}^\circ F_n \to f$ μ_L-a.e., for

$$\begin{aligned}\mu_L(\{{}^\circ F_n \not\to f\}) &\le \mu_L(\{{}^\circ F = f \wedge |f| < \infty\}^{\complement}) = \mu_L(\{{}^\circ F \ne f \vee |f| = \infty\}) \\ &\le \mu_L(\{{}^\circ F \ne f\}) + \mu_L(\{f = \infty\}) \\ &= 0\end{aligned}$$

Thus ${}^\circ F_n$ is Cauchy in the mean, since

$$\int |{}^\circ F_n - {}^\circ F_m| d\mu_L \le \int |{}^\circ F_n - f| d\mu_L + \int |{}^\circ F_m - f| d\mu_L,$$

and $\int |{}^\circ F_n \to f| d\mu_L \to 0$, by the dominated convergence theorem. (Note that ${}^\circ F_n$ are dominated by $|f|$.)

Consequently, the sequence F_n is S-Cauchy in the mean, since we have that

$$\int |F_n - F_m| d\mu = \int |{}^\circ F_n - {}^\circ F_m| d\mu_L,$$

by Proposition 2.3.11.

There must therefore exist an S-integrable S-limit in the mean F_κ, $\kappa \in \mathbb{N}_\infty$, for the $*$-sequence F_n, $n \in {}^*\mathbb{N}_{\ge 2}$, by Proposition 2.3.9.

Now note that $\mu_L(\{|{}^\circ F| = \infty\}) = 0$, since

$$\begin{aligned}\mu_L(\{|{}^\circ F| = \infty\}) &\le \mu_L(\{{}^\circ F = f = \infty\} \cup \{f \ne {}^\circ F = \infty\}) \\ &\le \mu_L(\{f = \infty\} \cup \{{}^\circ F \ne f\}) \\ &\le \mu_L(\{f = \infty\}) + \mu_L(\{{}^\circ F \ne f\}) \\ &= 0.\end{aligned}$$

Finally, we can conclude that the S-integrable F_κ is such that ${}^\circ F_\kappa = f$ μ_L-a.e., by observing that

$$\begin{aligned}\mu_L(\{{}^\circ F_\kappa \ne f\}) &= \mu_L(\{{}^\circ F_\kappa \ne f \wedge f = {}^\circ F\} \cup \{{}^\circ F_\kappa \ne f \wedge f \ne {}^\circ F\}) \\ &\le \mu_L(\{{}^\circ F_\kappa \ne {}^\circ F\} \cup \{f \ne {}^\circ F\}) \\ &= \mu_L(\{|F|^{-1}\left[[0,\kappa)\right]^{\complement}] \cup \{f \ne {}^\circ F\}) \\ &\le \mu_L(\{|F| = \infty\} \cup \{f \ne {}^\circ F\}) \\ &\le \mu_L(\{|F| = \infty\}) + \mu_L(\{f \ne {}^\circ F\}) \\ &= 0.\end{aligned}$$

□

Chapter 3

Anderson's Construction of Brownian Motion

In this section we will – as an application of the Loeb measure and integration theory we have developed – present the nonstandard construction of Browninan motion first devised by Anderson in his article *A non-standard representation for Brownian motion and Itô integration* (Anderson 1976).

We will, hopefully, convince the reader that this provides a more intuitive and/or simple treatment of a subject which is technically much more involved in standard probability theory.

3.1 Preliminaries

Recall the following definitions from probability theory.

Definition 3.1.1 (Probability Space)**.** A measure space $(\Omega, \mathcal{E}, p)$, where p is a measure of total measure 1 is called a *probability space*. The elements of the σ-algebra $\mathcal{E}$ are called *events* and p is called the probability measure.

Definition 3.1.2 (Random Variable)**.** A real valued function $X : \Omega \to \mathbb{R}$ measurable with respect to the σ-algebra $\mathcal{E}$ of a probability space $(\Omega, \mathcal{E}, p)$ is called a *random variable*.

Definition 3.1.3 (Stochastic Process)**.** A family of random variables $\{X_t \mid t \in T\}$ defined on a common probability space $(\Omega, \mathcal{E}, p)$ is called a *stochastic process*. The set T is called the *parameter set*.

For our arguments we shall also need Stirling's approximation.

Theorem 3.1.4 (Stirling's Approximation)**.** *For $n \in \mathbb{N}$,*

$$e^{\frac{1}{12n+1}} < \frac{n!e^n}{\sqrt{2\pi n}n^n} < e^{\frac{1}{12n}}.$$

Note in particular that for any $N \in \mathbb{N}_\infty$ we have

$$\frac{N!e^N}{\sqrt{2\pi N}N^N} = (1+\iota) \simeq 1, \tag{3.1}$$

where $\iota \simeq 0$.

Finally we shall need the nonstandard characterisations of the derivative and the Riemann integral. The proofs are completely mechanical. Recall the partial transfer function † and the partial transfer principle (1.1.29). In the converse of both we rely on enlargement for the existence of appropriate "infinitesimal" objects.

Proposition 3.1.5 (nonstandard characterisation of the derivative)**.** *A function $f : \mathbb{R} \to \mathbb{R}$ is differentiable at $x \in \mathbb{R}$ if and only if there exists a unique $f'(x) \in \mathbb{R}$ such that for any infinitesimal $\delta \simeq 0$ we have that*

$$f(x+\delta) - f(x) = f'(x)\delta \ + \delta\eta,$$

for some infinitesimal η.

Proof. Suppose f has derivative $f'(x)$ at x. Then,

$$\begin{aligned}
&\forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in \mathbb{R}_{>0} \forall \gamma \in [-\delta, \delta] \left(\left| \frac{f(x+\gamma) - f(x)}{\gamma} - f'(x) \right| < \varepsilon \right) \\
&\Rightarrow \forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in \mathbb{R}_{>0}{}^\dagger \left(\forall \gamma \in [-\delta, \delta] \left(\left| \frac{f(x+\gamma) - f(x)}{\gamma} - f'(x) \right| < \varepsilon \right) \right) \\
&\Rightarrow \forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in \mathbb{R}_{>0} \forall \gamma \in {}^*[-\delta, \delta] \left(\left| \frac{f(x+\gamma) - f(x)}{\gamma} - f'(x) \right| < \varepsilon \right) \\
&\Rightarrow \forall \varepsilon \in \mathbb{R}_{>0} \forall \gamma \in {}^*\mathbb{R}_{\simeq 0} \left(\left| \frac{f(x+\gamma) - f(x)}{\gamma} - f'(x) \right| < \varepsilon \right) \\
&\Rightarrow \forall \gamma \in {}^*\mathbb{R}_{\simeq 0} \left(\frac{f(x+\gamma) - f(x)}{\gamma} \simeq f'(x) \right) \\
&\Rightarrow \forall \gamma \in {}^*\mathbb{R}_{\simeq 0} \exists \eta \in {}^*\mathbb{R}_{\simeq 0} \left(\frac{f(x+\gamma) - f(x)}{\gamma} = f'(x) + \eta \right) \\
&\Rightarrow \forall \gamma \in {}^*\mathbb{R}_{\simeq 0} \exists \eta \in {}^*\mathbb{R}_{\simeq 0} \left(f(x+\gamma) - f(x) = f'(x)\gamma + \gamma\eta \right).
\end{aligned}$$

Conversely,

$$\begin{aligned}
&\forall \gamma \in {}^*\mathbb{R}_{\simeq 0} \exists \eta \in {}^*\mathbb{R}_{\simeq 0} (f(x+\gamma) - f(x) = f'(x)\gamma + \gamma\eta) \\
&\Rightarrow \forall \gamma \in {}^*\mathbb{R}_{\simeq 0} \exists \eta \in {}^*\mathbb{R}_{\simeq 0} \left(\frac{f(x+\gamma) - f(x)}{\gamma} = f'(x) + \eta \right) \\
&\Rightarrow \forall \gamma \in {}^*\mathbb{R}_{\simeq 0} \left(\frac{f(x+\gamma) - f(x)}{\gamma} \simeq f'(x) \right) \\
&\Rightarrow \exists \delta \in {}^*\mathbb{R}_{>0} \forall \gamma \in {}^*[-\delta, \delta] \left(\frac{f(x+\gamma) - f(x)}{\gamma} \simeq f'(x) \right) \\
&\Rightarrow \forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in {}^*\mathbb{R}_{>0} \forall \gamma \in {}^*[-\delta, \delta] \left(\left| \frac{f(x+\gamma) - f(x)}{\gamma} - f'(x) \right| < \varepsilon \right) \\
&\Rightarrow \forall \varepsilon \in \mathbb{R}_{>0}{}^{\dagger} \left(\exists \delta \in \mathbb{R}_{>0} \forall \gamma \in [-\delta, \delta] \left(\left| \frac{f(x+\gamma) - f(x)}{\gamma} - f'(x) \right| < \varepsilon \right) \right) \\
&\Rightarrow \forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in \mathbb{R}_{>0} \forall \gamma \in [-\delta, \delta] \left(\left| \frac{f(x+\gamma) - f(x)}{\gamma} - f'(x) \right| < \varepsilon \right).
\end{aligned}$$

□

To state the nonstandard characterisation of the Riemann integral we will need some terminology. We will only need the characterisation of the Riemann integral for functions of the form $f : [a, b] \to \mathbb{R}$. To this end a *partition* of $[a, b]$ refers to a finite collection of intervals that may be open, half-open, or closed the union of which is $[a, b]$. The diameter of an interval I with end points $a, b \in \mathbb{R}$ is denoted $d(I)$. Moreover, the diameter $d(P)$ of a partition P is the maximum of the diameters of its elements. We denote the collection of partitions of diameter $\delta \in \mathbb{R}_{>0}$ or less by Π_δ. A $*$-partition is infinitesimal if the diameter of the $*$-partition is infinitesimal. We denote the collection of infinitesimal $*$-partitions of ${}^*[a, b]$ by $\Pi_{\simeq 0}$. Finally, a choice is a function $\eta : P \to \bigcup P$ such that $\eta(p) \in p$ for each $p \in P$. We denote the set of choices on for a partition P by $C(P)$.

In these terms the nonstandard characterisation of the Riemann integrall is as follows:

Proposition 3.1.6 (nonstandard characterisation of the Riemann integral)**.** *A function $f : [a, b] \to \mathbb{R}$ has the Riemann integral $\int_a^b f(\xi)d\xi$ if and only if for every internal infinitesimal $*$-partition Π of ${}^*[a, b]$ and any choice $\eta : \Pi \to \bigcup \Pi$ we have that*

$$\int_a^b f(\xi)d\xi = {}^\circ \sum_p^{\Pi} p(\eta(p))d(p).$$

Proof. Suppose f is Riemann integrable with integral $\int_a^b f(\xi)d\xi = I$ over $[a,b]$.

$$\begin{aligned}
&\forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in \mathbb{R}_{>0} \forall P \in \Pi_\delta \forall \eta \in C(P) \left(\left| \sum_p^P f(\eta(p))d(p) - I \right| < \varepsilon \right) \\
&\Rightarrow \forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in \mathbb{R}_{>0}{}^\dagger \left(\forall P \in {}^*\Pi_\delta \forall \eta \in C(P) \left(\left| \sum_p^P f(\eta(p))d(p) - I \right| < \varepsilon \right) \right) \\
&\Rightarrow \forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in \mathbb{R}_{>0} \forall P \in {}^*\Pi_\delta \forall \eta \in {}^*C(P) \left(\left| \sum_p^P f(\eta(p))d(p) - I \right| < \varepsilon \right) \\
&\Rightarrow \forall \varepsilon \in \mathbb{R}_{>0} \forall P \in {}^*\Pi_{\simeq 0} \forall \eta \in {}^*C(P) \left(\left| \sum_p^P f(\eta(p))d(p) - I \right| < \varepsilon \right) \\
&\Rightarrow \forall P \in {}^*\Pi_{\simeq 0} \forall \eta \in {}^*C(P) \left(\left| \sum_p^P f(\eta(p))d(p) - I \right| \simeq 0 \right) \\
&\Rightarrow \forall P \in {}^*\Pi_{\simeq 0} \forall \eta \in {}^*C(P) \left(\sum_p^P f(\eta(p))d(p) \simeq I \right).
\end{aligned}$$

Conversely,

$$\begin{aligned}
&\forall P \in {}^*\Pi_{\simeq 0} \forall \eta \in {}^*C(P) \left(\sum_p^P f(\eta(p))d(p) \simeq I \right) \\
&\Rightarrow \exists \delta \in {}^*\mathbb{R}_{>0} \forall P \in {}^*\Pi_\delta \forall \eta \in {}^*C(P) \left(\sum_p^P f(\eta(p))d(p) \simeq I \right) \\
&\Rightarrow \forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in {}^*\mathbb{R}_{>0} \forall P \in {}^*\Pi_\delta \forall \eta \in {}^*C(P) \left(\left| \sum_p^P f(\eta(p))d(p) - I \right| < \varepsilon \right) \\
&\Rightarrow \forall \varepsilon \in \mathbb{R}_{>0}{}^\dagger \left(\exists \delta \in \mathbb{R}_{>0} \forall P \in \Pi_\delta \forall \eta \in C(P) \left(\left| \sum_p^P f(\eta(p))d(p) - I \right| < \varepsilon \right) \right) \\
&\Rightarrow \forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in \mathbb{R}_{>0} \forall P \in \Pi_\delta \forall \eta \in C(P) \left(\left| \sum_p^P f(\eta(p))d(p) - I \right| < \varepsilon \right).
\end{aligned}$$

□

3.2 Anderson's Brownian Motion

The idea behind Anderson's construction of one dimensional Brownian motion is essentially to formalise the intuitive picture. Namely, that of an infinitely fast random walk of infinitesimal steps. To this end let $N \in \mathbb{N}_\infty$, let $\delta = \frac{1}{N}$, and

$$T = \{n\delta \mid n \in {}^*\mathbb{N}_{\leq N}\}.$$

We will refer to T as the *hyperfinite timeline*. The elements in T which are even multiples of δ are called the even time steps, similarly the odd multiples are called the odd time steps. Now, the internal walks make up the $*$-finite set $\Omega = \{-1,1\}^T$. On which we can define the internal probability charge by:

$$P(A) = \frac{|A|}{2^N},$$

where A is any internal subset of Ω. Since, there are 2^N internal walks in total we have $P(\Omega) = 1$. It is simple to check that P is an internal probability charge defined on the internal algebra

$\mathcal{A}$ of internal subsets of Ω, that is $\mathcal{A} = \mathcal{P}_T(\Omega)$. The fact that every internal subset of Ω is P-measurable will turn out to be extremely useful. In fact this is a feature of many nonstandard constructions and lies at the heart of why hyperfinite Loeb measures are such a powerful tool. We will see this first hand when we construct the Wiener measure later.

We can now define Anderson's Random Walk

Definition 3.2.1 (Anderson's Random Walk)**.** For each $t \in T$ and $\omega \in \Omega$ let

$$B_t(\omega) = \sum_s^{T_{<t}} \omega(s)\sqrt{\delta}.$$

The $\sqrt{\delta}$ term is chosen so as to normalise the variance of B_t. The idea is that the size of the spatial step is somehow coupled to the size of the time step. Suppose that the step size is $f(\delta)$ for any choice of time step δ. We can now solve for the $f(\delta)$ which normalises the variance.

$$\begin{aligned} 1 = \mathrm{Var}(\sum_s^{T_{<1}} \omega(s)f(\delta)) &= \mathbb{E}\left(\left(\sum_s^{T_{<1}} \omega(s)f(\delta)\right)^2\right) - \mathbb{E}\left(\left(\sum_s^{T_{<1}} \omega(s)f(\delta)\right)\right)^2 \\ &= \sum_s^{T_{<1}} \mathbb{E}\left(\omega(s)^2 f(\delta)^2\right) = \sum_s^{T_{<1}} \mathbb{E}(f(\delta)^2) \\ &= N\mathbb{E}f(\delta)^2 = Nf(\delta)^2. \end{aligned}$$

Thus,

$$\sqrt{\delta} = \sqrt{\frac{1}{N}} = f(\delta).$$

Note that each B_t is measurable with respect to $\mathcal{A}$. We will spend most of the rest of this section proving that ${}^\circ B_{\lfloor t \rfloor_\delta} : [0,1] \times \Omega \to \mathbb{R}$ is Brownian process in the standard sense. (Recall that $\lfloor x \rfloor_\delta$ stands for the greatest multiple of δ smaller than x.) We therefore make the following definition.

Definition 3.2.2 (Anderson's Brownian Process)**.**

$$b_t = {}^\circ B_{\lfloor t \rfloor_\delta}, \quad t \in [0,1],$$

Note that each b_t is measurable with respect to the Loeb measure P_L, by the uniform lifting lemma (2.2.7), and $\{b_t\}_t^{[0,1]}$ is a stochastic process since the Loeb measure P_L is a probability measure and the Loeb space $(P_L, \mathcal{A}_L, \Omega)$ is a probability space. Moreover, it turns out that b_t is not dependent on the choice of infinitesimal δ. (Much as our construction of the Lebesgue measure in 2.1.1 The Lebesgue Measure did not depend on the choice of infinitesimal.) If our presentation for some reason doesn't convince the concerned reader see the original paper (Anderson 1976). In it Anderson places more emphasis on convincing the reader of this fact.

As we already stated, our goal now is to show that $b_t, t \in [0,1]$ is a Brownian process. Recall, that a Brownian process is defined as follows:

Definition 3.2.3 (Brownian Process)**.** A stochastic process $b_t(\omega)$ is called a *Brownian process* if it satisifies the following criteria:

1. $b_t(\cdot)$ is measurable for each $t \in [0,1]$.
2. for any real numbers $s, t \in [0,1]$ such that $t > s$ the random variable $b_t(\cdot) - b_s(\cdot)$ is normally distributed with mean 0 and variance $t - s$.

3. for any choice $s_0 \le t_0 \le s_1 \le t_1 \le \cdots \le s_{n-1} \le t_{n-1}$ of $n \in \mathbb{N}$ real numbers the random variables $b_{t_i}(\cdot) - b_{s_i}(\cdot)$ are independent.

Based on our earlier observations we immediately note that.

Proposition 3.2.4. *The stochastic process b_t, $t \in [0,1]$ satisfies the first defining criterion of a brownian process.*

To prove the third criterion we shall need the following nonstandard form of De Moivre's Limit Theorem for Anderson's Random Walk.

Theorem 3.2.5 (De Moivre's Limit Theorem)**.** *If $x \in {}^*\mathbb{R}$ is finite and $t \in T$ isn't infinitesimal, then*

$$P(x < B_t(\omega) \le x + \delta') = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} \delta'(1+\iota),$$

where B_t is Anderson's random walk, $\delta' = 2\sqrt{\delta}$, and ι is some infinitesimal. Moreover, if a, b are finite $*$-reals and $t \in T$ such that $t \not\simeq 0$, then

$$P(a < B_t(\omega) \le b) \simeq \frac{1}{\sqrt{2\pi t}} \int_a^b e^{-\frac{\xi^2}{2t}} d\xi.$$

In other words, $B(t)$ is nearly normally distributed with mean 0 and variance t.

The proof largely follows the presentation in *Foundations of Infinitesimal Stochastic Analysis* (Stroyan and Bayod 1986).

Proof. We begin by deriving an expression for the probability that we have not moved at time some even time step $t \in T$. Since each $t \in T$ is an even multiple of δ we have $t = \frac{2n}{N} = 2n\delta$. Then,

$$P(B_t(\omega) = 0) = \frac{1}{2^{2n}} \binom{2n}{n} = \frac{(2n)!}{2^{2n} n! n!},$$

by the binomial theorem,

$$\begin{aligned} &= \frac{(2n)^{2n} \sqrt{2\pi 2n} (e^n)^2}{2^{2n} (n^n)^2 (\sqrt{2\pi n})^2 (e^2 n)} (1+\iota_t) = \frac{1}{\sqrt{\pi n}} (1+\iota_t) \\ &= \frac{1}{\sqrt{\pi t/(2\delta)}} (1+\iota_t) = \frac{2}{\sqrt{2\pi t}} \sqrt{\delta}(1+\iota_t) = \frac{\delta'}{\sqrt{2\pi t}} (1+\iota_t) \end{aligned}$$

Next, we derive an expression for the probability that we have moved a finite non-infinitesimal distance x, such that $|x| = k2\sqrt{\delta} = k\delta'$, by an non-infinitesimal even time step t. Observe that at an even time step we must have moved by an even multiple of $\sqrt{\delta}$. For, we have made an even number of $\pm\sqrt{\delta}$ sized steps. That is why we choose $\delta' = 2\sqrt{\delta}$. Note that k/n is an infinitesimal, since

$$\frac{k}{n} = \frac{k\delta'}{n\delta'} = \frac{|x|}{t}\sqrt{\delta},$$

and $\sqrt{\delta} \simeq 0$. Moreover, $\frac{k(k-1)}{n} = (1+\tau)\frac{x^2}{t}$ for an infinitesimal τ, since

$$\frac{k(k-1)}{n} = \frac{k^2}{n} + \frac{k}{n} = \frac{|x|^2}{2t} + \frac{|x|^2}{2t}\frac{2\sqrt{\delta}}{|x|} = \frac{x^2}{2t}(1+\tau).$$

That is $\tau = \frac{\delta'}{|x|}$. Now,

$$\begin{aligned}
P(B_t(\omega) = x) &= \frac{1}{2^{2n}} \binom{2n}{n+k} \\
&= \frac{1}{2^{2n}} \binom{2n}{n} \frac{\prod_h^k (n-h)}{\prod_h^{(k+1)_{>0}} h} \\
&= \frac{\delta'}{\sqrt{2\pi t}} (1+\iota_t) \frac{n}{n+k} \prod_h^{k_{>0}} \frac{1-\frac{h}{n}}{1+\frac{h}{n}} \\
&= \frac{\delta'}{\sqrt{2\pi t}} (1+\iota_t) \frac{n}{n+k} \prod_h^{k_{>0}} \left(1 - \frac{2h}{n} + \frac{2h}{n}\frac{h}{n+h}\right) \\
&= \frac{\delta'}{\sqrt{2\pi t}} (1+\iota_t) \frac{1}{1+\frac{k}{n}} \exp\left(\sum_h^{k_{>0}} \ln\left(1 - \frac{2h}{n}(1+\iota_h)\right)\right),
\end{aligned}$$

where $\iota_h = \frac{h}{n+h}$ is an infinitesimal, since $\frac{h}{n+h} < \frac{k}{n} \simeq 0$

$$P(B_t(\omega) = x) == \frac{\delta'}{\sqrt{2\pi t}} (1+\iota_t) \frac{1}{1+\frac{k}{n}} \exp\left(\sum_h^{k_{>0}} \left(-\frac{2h}{n}(1+\theta_h)\right)\right)$$

where $\theta_h = \iota_h + \lambda_h$ for an infinitesimal λ_h such that

$$\ln(1 - \frac{2h}{n}(1+\iota_h)) = \ln(1 - \frac{2h}{n}(1+\iota_h)) - \ln(1) = -\frac{2h}{n}(1+\iota_h+\lambda_h).$$

Such a λ exists, by the nonstandard characterisation of the derivative (3.1.5), since ln is differentiable at 1 and $\frac{2h}{n}$ is infinitesimal.

$$P(B_t(\omega) = x) == \frac{\delta'}{\sqrt{2\pi t}} (1+\iota_t) \frac{1}{1+\frac{k}{n}} \exp\left(\frac{-k(k-1)}{n}(1+\theta)\right),$$

where θ is some inifinitesimal,

$$P(B_t(\omega) = x) == \frac{\delta'}{\sqrt{2\pi t}} (1+\iota_t) \frac{1}{1+\frac{k}{n}} \exp\left(-\frac{x^2}{t}(1+\rho)\right),$$

where $\rho = (\theta + \tau + \theta\tau)$, since $\frac{k(k-1)}{n}(1+\theta) = \frac{x^2}{t}(1+\tau)(1+\theta)$.

$$P(B_t(\omega) = x) == \frac{\delta'}{\sqrt{2\pi t}} \exp\left(\frac{-x^2}{2t}\right)(1+\iota),$$

where ι is the infinitesimal such that $(1+\iota) = (1+\frac{k}{n})^{-1}(1+\iota_t)(1+\iota')$, for $\iota \simeq 0$ such that $(1+\iota') = e^{\rho\frac{x^2}{n}}$.

In the case where t is an odd non-infinitesimal time step and we have moved by a finite non-infinitesimal distance $|x|$, $x \in {}^*\mathbb{R}$. Let n be such that $(2n+1)\delta = t$, let k be such that $(2k+1)\sqrt{\delta} = |x|$, and $\delta' = 2\sqrt{\delta}$. Observe that at an odd time step we must have moved by an odd multiple of $\sqrt{\delta'}$. The argument now follows very similarly to the even case. In particular, $\frac{k}{n}$ is infinitesimal, $k(k-1) = (1+\tau)\frac{x^2}{t}$ for some infinitesimal τ, and, moreover,

$$\frac{1}{2^{2n}} \binom{2n}{n+k} = \frac{\delta'}{\sqrt{2\pi t}}(1+\iota_t),$$

for some infinitesimal ι_t. Now,

$$P(B_t(\omega)=x)=\frac{1}{2^{2n+1}}\binom{2n+1}{n+k}$$
$$=\frac{2n+1}{2(n+1-k)}\frac{1}{2^{2n}}\binom{2n}{n+k}$$
$$=(1+\varepsilon)\frac{1}{2^{2n}}\binom{2n}{n+k}$$

where ε is an infinitesimal such that $(1+\varepsilon)=\frac{2n+1}{2(n+1-k)}$, such an ε exists since $\frac{2n+1}{2(n+1-k)}\simeq 1$.

$$=(1+\varepsilon)\frac{\delta'}{\sqrt{2\pi t}}\exp\left(\frac{-x^2}{2t}\right)(1+\iota),$$

by a similar line of reasoning as in the even case. This concludes the proof of the first part of the theorem.

The second part is a nearly immediate consequence of the first part of the theorem and the nonstandard characterisation of the Riemann integral (3.1.6). □

We can now prove that b_t satisfies the second defining criterion of a Brownian process.

Proposition 3.2.6. *The stochastic process b_t, $t\in[0,1]$ satisfies the second defining criterion of a Brownian process.*

Proof. Let $t,s\in T$ and $t>s$. Then,

$$P(\{\omega\in\{-1,1\}^T | B_t(\omega)-B_s(\omega)\in(a,b]\})=P\left(\left\{\omega\in\{-1,1\}^T\middle|\sum_h^{T_{\le t}\setminus T_{\le s}}\omega(h)\sqrt{\delta}\in(a,b]\right\}\right)$$
$$=P\left(\left\{\omega\in\{-1,1\}^T\mid\sum_h^{T_{\le t-s}}\omega(h)\sqrt{\delta}\in(a,b]\right\}\right)=P(\{\omega\in\{-1,1\}^T\mid B_{t-s}\in(a,b]\})$$
$$\simeq\frac{1}{\sqrt{2\pi(t-s)}}\int_a^b e^{-\frac{\xi^2}{2(t-s)}}d\xi,$$

by De Moivre's Limit Theorem (3.2.5).

Now let t',s' be such that $t=\lfloor t'\rfloor_\delta$ and $s=\lfloor s'\rfloor_\delta$

$$P_L(b_{t'}(\omega)-b_{s'}(\omega)\in(a,b])=P_L\left(\bigcup_m^{\mathbb{N}}\bigcap_n^{\mathbb{N}}(B_t-B_s)^{-1}\left[\left(a+\frac{1}{m},b+\frac{1}{n}\right]\right]\right)$$
$$=\lim_{m\to\infty}\lim_{n\to\infty}P_L\left((B_t-B_s)^{-1}\left[\left(a+\frac{1}{m},b+\frac{1}{n}\right]\right]\right)$$
$$=\lim_{m\to\infty}\lim_{n\to\infty}{}^\circ\left(\frac{1}{\sqrt{2\pi(t'-s')}}\int_{a+\frac{1}{m}}^{b+\frac{1}{n}}e^{-\frac{\xi^2}{2(t'-s')}}d\xi\right)$$
$$=\lim_{m\to\infty}\lim_{n\to\infty}{}^\circ\left(\frac{1}{\sqrt{2\pi(t'-s')}}\int_{a+\frac{1}{m}}^{b+\frac{1}{n}}e^{-\frac{\xi^2}{2(t'-s')}}d\xi\right)$$
$$=\lim_{m\to\infty}\lim_{n\to\infty}\frac{1}{\sqrt{2\pi(t'-s')}}\int_{a+\frac{1}{m}}^{b+\frac{1}{n}}e^{-\frac{\xi^2}{2(t'-s')}}d\xi$$
$$=\frac{1}{\sqrt{2\pi(t'-s')}}\int_a^b e^{-\frac{\xi^2}{2(t'-s')}}d\xi.$$

□

The proof that b_t, $t \in [0,1]$ satisfies the third criterion is straighforward but tedious. We begin by proving the following proposition we shall also need later.

Proposition 3.2.7. *For any* $0 \leq s_0 \leq t_0 \leq s_1 \leq \cdots \leq s_{n-1} \leq t_{n-1} \leq 1$ *where* $n \in {}^*\mathbb{N}$ *and* $t_i, s_i \in T$ *the random variables* $B_{t_i} - B_{s_i}$ *are* $*$*-independent.*

Proof. Let $\{s_i\}_i^n$, $\{t_i\}_i^n$, and $n \in {}^*\mathbb{N}$ be as in the statement of the proposition. Furthermore, let $(a_i, b_i] \subseteq {}^*\mathbb{R}$ for each $i \in n$. Moreover, let $r = \{r \in T \mid \exists i \in n(t_i < r \leq s_{i+1})\}$ and $T_{t_i,s_i} = T \cap {}^*(s_i, t_i]$ and $H = \{-1, 1\}$. Now,

$$P\left(\bigcap_i^n (B_{t_i} - B_{s_i})^{-1}\left[(a_i, b_i]\right]\right) = P\Big(\Big\{\omega \in H^T \Big| \forall i \in n(B_{t_i}(\omega) - B_{s_i}(\omega) \in (a_i, b_i])\Big\}\Big)$$

$$= P\left(\left\{\omega \in H^T \middle| \forall i \in n\left(\sum_h^{T_{t_i,s_i}} \omega(h)\sqrt{\delta} \in (a_i, b_i]\right)\right\}\right)$$

$$= P\left(\left\{\omega \in H^r \times \prod_i^n H^{T_{t_i,s_i}} \middle| \forall i \in n\left(\sum_h^{T_{t_i,s_i}} \omega(h)\sqrt{\delta} \in (a_i, b_i]\right)\right\}\right),$$

since $r \cup \bigcup_i^n T_{t_i,s_i} = T$ and $r, T_{t_0,s_0}, \ldots, T_{t_{n-1},s_{n-1}}$ are disjoint. Moreover,

$$P\left(\bigcap_i^n (B_{t_i} - B_{s_i})^{-1}\left[(a_i, b_i]\right]\right) = P\left(H^r \times \prod_i^n \left\{\omega \in H^{T_{t_i,s_i}} \middle| \sum_h^{T_{t_i,s_i}} \omega(h)\sqrt{\delta} \in (a_i, b_i]\right\}\right)$$

$$= \frac{\left|H^r \times \prod_i^n \left\{\omega \in H^{T_{t_i,s_i}} \middle| \sum_h^{T_{t_i,s_i}} \omega(h)\sqrt{\delta} \in (a_i, b_i]\right\}\right|}{2^{|T|}}$$

$$= \frac{|H^r| \times \prod_i^n \left|\left\{\omega \in H^{T_{t_i,s_i}} \middle| \sum_h^{T_{t_i,s_i}} \omega(h)\sqrt{\delta} \in (a_i, b_i]\right\}\right|}{2^{|r|} \prod_i^n 2^{t_i - s_i}}$$

$$= \frac{|H^r|}{2^{|r|}} \times \prod_i^n \frac{\left|\left\{\omega \in H^{T_{t_i,s_i}} \middle| \sum_h^{T_{t_i,s_i}} \omega(h)\sqrt{\delta} \in (a_i, b_i]\right\}\right|}{2^{t_i - s_i}}$$

$$= \prod_i^n \frac{\left|\left\{\omega \in H^{T_{t_i,s_i}} \middle| \sum_h^{T_{t_i,s_i}} \omega(h)\sqrt{\delta} \in (a_i, b_i]\right\}\right| 2^{|T| - (t_i - s_i)}}{2^{|T|}}$$

since $|H^r| = 2^{|r|}$ and $T_{t_i - s_i}$ is a strict subset of T with $t_i - s_i$ elements for each $i \in n$. Finally,

$$P\left(\bigcap_i^n (B_{t_i} - B_{s_i})^{-1}\left[(a_i, b_i]\right]\right) = \prod_i^n \frac{\left|\left\{\omega \in H^{T_{t_i,s_i}} \middle| \sum_h^{T_{t_i,s_i}} \omega(h)\sqrt{\delta} \in (a_i, b_i]\right\}\right| 2^{|T \setminus T_{t_i,s_i}|}}{2^{|T|}}$$

$$= \prod_i^n \frac{\left|\left\{\omega \in H^{T_{t_i,s_i}} \middle| \sum_h^{T_{t_i,s_i}} \omega(h)\sqrt{\delta} \in (a_i, b_i]\right\}\right| \left|H^{T \setminus T_{t_i,s_i}}\right|}{2^{|T|}}$$

$$= \prod_i^n \frac{\left|\left\{\omega \in H^{T_{t_i,s_i}} \times H^{T \setminus T_{t_i,s_i}} \middle| \sum_h^{T_{t_i,s_i}} \omega(h)\sqrt{\delta} \in (a_i, b_i]\right\}\right|}{2^{|T|}}$$

$$= \prod_i^n \frac{\left|\left\{\omega \in H^{T} \middle| \sum_h^{T_{t_i,s_i}} \omega(h)\sqrt{\delta} \in (a_i, b_i]\right\}\right|}{2^{|T|}}$$

$$= \prod_i^n P\left(\left\{\omega \in H^T \middle| \sum_h^{T_{t_i,s_i}} \omega(h)\sqrt{\delta} \in (a_i, b_i]\right\}\right)$$

$$= \prod_i^n P\left(\left\{\omega \in H^T \middle| \omega \in (B_{t_i} - B_{s_i})^{-1}[(a_i, b_i]]\right\}\right).$$

□

Proposition 3.2.8. *The stochastic process b_t, $t \in [0,1]$ satisfies the third defining criterion of a Brownian process.*

Proof. Let $0 \leq s'_0 < t'_0 \leq s'_1 < t'_1 \leq \cdots \leq s'_{n-1} < t'_{n-1} \leq 1$ where $n \in \mathbb{N}$ and let $(a_i, b] \subseteq \mathbb{R}$, $i \in n$. Moreover, let $t_i = \lfloor t'_i \rfloor_\delta$ and $s_i = \lfloor s'_i \rfloor_\delta$ for each $i \in n$. Then,

$$P_L\left(\bigcap_i^n (b_{t'_i} - b_{s'_i})^{-1}[(a_i, b_i]]\right) = P_L\left(\bigcup_j^{\mathbb{N}} \bigcap_k^{\mathbb{N}} \bigcap_i^n (B_{t_i} - B_{s_i})^{-1}\left[{}^*\left(a_i + \frac{1}{j}, b_i + \frac{1}{k}\right]\right]\right)$$

$$= \lim_{j\to\infty} \lim_{k\to\infty} P_L\left(\bigcap_i^n (B_{t_i} - B_{s_i})^{-1}\left[{}^*\left(a_i + \frac{1}{j}, b_i + \frac{1}{k}\right]\right]\right)$$

$$= \lim_{j\to\infty} \lim_{k\to\infty} {}^\circ P\left(\bigcap_i^n (B_{t_i} - B_{s_i})^{-1}\left[{}^*\left(a_i + \frac{1}{j}, b_i + \frac{1}{k}\right]\right]\right)$$

$$= \lim_{j\to\infty} \lim_{k\to\infty} {}^\circ \prod_i^n P\left((B_{t_i} - B_{s_i})^{-1}\left[{}^*\left(a_i + \frac{1}{j}, b_i + \frac{1}{k}\right]\right]\right),$$

by Proposition 3.2.7. Moreover,

$$P_L\left(\bigcap_i^n (b_{t'_i} - b_{s'_i})^{-1}[(a_i, b_i]]\right) = \lim_{j\to\infty} \lim_{k\to\infty} \prod_i^n {}^\circ P\left((B_{t_i} - B_{s_i})^{-1}\left[{}^*\left(a_i + \frac{1}{j}, b_i + \frac{1}{k}\right]\right]\right)$$

$$= \lim_{j\to\infty} \lim_{k\to\infty} \prod_i^n P_L\left((B_{t_i} - B_{s_i})^{-1}\left[{}^*\left(a_i + \frac{1}{j}, b_i + \frac{1}{k}\right]\right]\right)$$

$$= \prod_i^n \lim_{j\to\infty} \lim_{k\to\infty} \prod_i^n P_L\left((B_{t_i} - B_{s_i})^{-1}\left[{}^*\left(a_i + \frac{1}{j}, b_i + \frac{1}{k}\right]\right]\right)$$

$$= \prod_i^n P_L\left(\bigcup_j^{\mathbb{N}} \bigcap_k^{\mathbb{N}} (B_{t_i} - B_{s_i})^{-1}\left[{}^*\left(a_i + \frac{1}{j}, b_i + \frac{1}{k}\right]\right]\right)$$

$$= \prod_i^n P_L((b_{t'_i} - b_{s'_i})^{-1}[(a, b]]).$$

□

We have thus shown that

Theorem 3.2.9. *The stochastic process* b_t, $t \in [0,1]$ *is a Brownian process.*

Proof. It satisfies the first defining criterion by Proposition 3.2.4, the second defining criterion by Proposition 3.2.6, and the third defining criterion by Proposition 3.2.8. □

3.3 The Wiener Measure

Anderson's Brownian motion can also be used to give an elegant construction of the Wiener measure. We adopt the strategy presented in *Nonstandard Methods in Stochastic Analysis and Mathematical Physics* (Albeverio et al. 1986). Let $C[0,1]$ stand for the Banach space of real valued continuous functions with domain $[0,1]$ equipped with the sup-norm.

Recall that the Wiener measure is defined:

Definition 3.3.1 (Wiener Measure)**.** The *Wiener measure* is defined as the unique Borel measure on $C[0,1]$ that satisfies the following conditions:

1. The random variable $f_t - f_s : C[0,1] \to \mathbb{R}; g \mapsto g(t) - g(s)$ is normally distributed with mean 0 and variance $t - s$ for each $t, s \in [0,1]$ such that $t > s$
2. For each choice $t_i, s_i \in [0,1]$, $i \in n$ of $n \in \mathbb{N}$ real numbers such that $s_i < t_i$ for each $i \in n$ and $t_i \leq s_{i+i}$ for each $i \in (n-1)$ the random variables $f_{t_i} - f_{s_i}$ are independent.

For our construction of the Wiener measure we need the following nonstandard notion.

Definition 3.3.2 (S-continuity)**.** A function $f : {}^*\mathbb{R} \to {}^*\mathbb{R}$ is called *S-continuous* if for each $x \in \mathbb{R}$ and $x' \in {}^*\mathbb{R}$ such that $x \simeq x'$ we have that

$$f(x) \simeq f'(x).$$

One motivation for this definition is the following.

Proposition 3.3.3 (nonstandard characterisation of continuity)**.** *A function* $f : X \to Y$ *is continuous precisely when* ${}^*f$ *is S-continuous.*

Moreover, the following property of S-continuous functions is of interest to us.

Proposition 3.3.4. *If* $f : {}^*\mathbb{R} \to {}^*\mathbb{R}$ *is S-continuous, then* ${}^\circ f : \mathbb{R} \to \mathbb{R}$ *is continuous.*

Proof. Suppose that f is S-continuous. It follows that

$$\forall x \in \mathbb{R} \forall \varepsilon \in \mathbb{R}_{>0} \forall \delta \in {}^*\mathbb{R}_{\simeq 0} \forall x' \in {}^*\mathbb{R}(|x - x'| < \delta \to |f(x) - f(x')| < \varepsilon),$$

by the definition of $\simeq$. Moreover,

$$\forall x \in \mathbb{R}\, \forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in \mathbb{R}_{>0} \forall x' \in {}^*\mathbb{R}(|x - x'| < \delta \to \ |f(x) - f(x')| < \varepsilon).$$

For, since the internal proposition

$$\forall x' \in {}^*\mathbb{R}(|x - x'| < \delta \to \ |f(x) - f(x')| < \varepsilon,$$

is true for every infinitesimal δ, it follows that it must hold for some $\delta \ \in \mathbb{R}_{>0}$ by Cauchy's principle (1.5.2) for each $x \in \mathbb{R}$ and $\varepsilon \in \mathbb{R}_{>0}$. Furthermore,

$$\forall x \in \mathbb{R} \forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in \mathbb{R}_{>0} \forall x' \in \mathbb{R}(|x - x'| < \delta \to \ |f(x) - f(x')| < \varepsilon),$$

since $\mathbb{R} \subset {}^*\mathbb{R}$. Finally, we have that

$$\forall x \in \mathbb{R} \forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in \mathbb{R}_{>0} \forall x' \in \mathbb{R}(|x - x'| < \delta \to |{}^\circ f(x) - {}^\circ f(x')| < 2\varepsilon),$$

since

$$|{}^\circ f(x) - {}^\circ f(x')| \leq |{}^\circ f(x) - f(x)| + |f(x) - f(x')| + |f(x') - {}^\circ f(x')| \lesssim \varepsilon < 2\varepsilon,$$

for any $\varepsilon \in \mathbb{R}_{>0}$ when $|f(x) - f(x')| < \varepsilon$. □

Let $SC[0,1]$ stand for the collection of S-continuous functions on ${}^*[0,1]$. The notion of S-continuity is related to the monads of $C[0,1]$ as follows

Proposition 3.3.5. *Suppose that $f \in C[0,1]$. Then*

$$m(f) = \{g \in SC[0,1] \cap {}^*C[0,1] \mid \forall x \in [0,1](g(x) \simeq f(x))\}.$$

Proof. For any $f \in C[0,1]$ the corresponding monad is, by definition,

$$m(f) = \{g \in {}^*C[0,1] \mid {}^*\|f - g\| \simeq 0\}.$$

Now, we have the following sequence of equivalences:

$$\begin{aligned} &{}^*\|f - g\| \simeq 0 \Leftrightarrow \forall x \in {}^*[0,1](f(x) \simeq g(x)) \\ &\Leftrightarrow \forall x \in [0,1](f(x) \simeq g(x) \wedge \forall x' \in {}^*[0,1](x \simeq x' \to \ g(x) \simeq f(x) \simeq f(x') \simeq g(x')) \\ &\Leftrightarrow \forall x \in [0,1](f(x) \simeq g(x)) \wedge g \in SC[0,1]. \end{aligned}$$

It follows that

$$m(f) = \{g \in SC[0,1] \cap {}^*C[0,1] \mid \forall x \in [0,1](g(x) \simeq f(x))\}.$$

□

The expected value with respect to P has the following unsurprising definition. (Technically this is the $*$-expected value but we shall refer to it as the expected value.)

Definition 3.3.6 (Expected Value)**.** Let X be a random variable with respect to an internal probability charge P on the event space Ω. Then the *expected value* if it exists is defined

$$\mathbb{E}_P(X) = \int_\Omega X dP.$$

In particular, when X is a $*$-simple function – which is always the case when the Loeb measure is hyperfinite – the integral takes the form of a $*$-sum. Namely,

$$\mathbb{E}_P(X) = \sum_{v}^{X^{-1}[\mathrm{ran}(X)]} vP(X^{-1}[\{v\}]).$$

We will need the Markov inequality

Proposition 3.3.7 (Markov's Inequality)**.** *Suppose $f : \mathbb{R} \to \mathbb{R}$ is a continuous non-negative function, $X : \Omega \to \mathbb{R}$ is a random variable on a probability space $(\Omega, \mathcal{A}, P)$, and $\lambda \in \mathbb{R}$ a non-negative real number, then*

$$P(f(X) \leq \lambda) \leq \frac{\mathbb{E}(f(X))}{\lambda}.$$

Proof. Let f, X, $(\Omega, \mathcal{A}, P)$, and λ be as in the statement of the proposition. It follows that $A = (f \circ X)^{-1}[[\lambda, \infty)]$ is measurable since $f \circ X$ is measurable. Now,

$$\begin{aligned}\mathbb{E}(f(X)) &= \int_\Omega f(X)dP \geq \int_A f(X)dP \geq \int_A \lambda dP \\ &= \lambda \int_A dP = \lambda P(f(X) < \lambda).\end{aligned}$$

□

We can now prove the following lemma for Anderson's Random Walk (3.2.1) B_t.

Lemma 3.3.8. *If $t \in T$, then*

$$\mathbb{E}(B_t^4) \leq 3t^2.$$

Note that B_t is Anderson's random walk.

Proof. Suppose that $t \in T$, and let $\Delta_x = \omega(x)\sqrt{\delta}$ for each $x \in T$. Equivalently, let $\Delta_x = B_{x+\delta} - B_x$ for each $x \in T$. We begin by observing that,

$$B_t^n = \sum_r^{T_{<t}} (B_{t+\delta}^n - B_t^n),$$

for each $n \in \mathbb{N}$ and $t \in T$, since $B_0 = 0$.

Next we prove that $\mathbb{E}(B_t^2) = t$. We proceed by a direct calculation. We know that

$$\mathbb{E}(B_t^2) = \mathbb{E}\left(\sum_r^{T_{<t}} (B_{r+\delta}^2 - B_r^2)\right),$$

by applying the equality we proved in the first part. Moreover,

$$\mathbb{E}(B_t^2) = \sum_r^{T_{<t}} \mathbb{E}\left(B_{r+\delta}^2 - B_r^2)\right),$$

since $\mathbb{E}$ is $*$-additive by transfer. Furthermore,

$$\begin{aligned}\mathbb{E}(B_t^2) &= \sum_r^{T_{<t}} \mathbb{E}\left((\Delta_r + B_r)^2 - B_r^2)\right) \\ &= \sum_r^{T_{<t}} \mathbb{E}\left(2\Delta_r B_r + \Delta_r^2\right) \\ &= \sum_r^{T_{<t}} 2\mathbb{E}(\Delta_r B_r) + \mathbb{E}(\Delta_r^2)\end{aligned}$$

since $\mathbb{E}$ is $*$-linear by transfer. Moreover,

$$\mathbb{E}(B_t^2) = \sum_r^{T_{<t}} 2\mathbb{E}(\Delta_r)\mathbb{E}(B_r) + \mathbb{E}(\Delta_r^2)$$

since $\Delta_r = B_{r+\delta} - B_r$ and B_r are $*$-independent, by Proposition 3.2.7. Now,

$$\mathbb{E}(B_t^2) = \sum_r^{T_{<t}} \delta,$$

since $\mathbb{E}(\Delta_r^2) = \delta$ and $\mathbb{E}(\Delta_r) = 0$. Finally,

$$\mathbb{E}(B_t^2) = \frac{t}{\delta}\delta = t.$$

Finally we prove the inequality $\mathbb{E}(B_t^4) \leq 3t^2$. The arguments are essentially the same as in the above proof that $\mathbb{E}(B_t^2) = t$. Because of this we proceed in a more succinct manner, elaborating only on the steps where we make use of arguments we have not yet employed.

$$\begin{aligned}
\mathbb{E}(B_t^4) &= \mathbb{E}\left(\sum_r^{T_{<t}} B_{r+\delta}^4 - B_r^4)\right) \\
&= \sum_r^{T_{<t}} \mathbb{E}((\Delta_r + B_r)^4 - B_r^4) \\
&= \sum_r^{T_{<t}} \left(\mathbb{E}(\Delta_r^4) + 4\mathbb{E}(\Delta_r)\mathbb{E}(B_r^3) + 6\mathbb{E}(\Delta_r^2)\mathbb{E}(B_r^2) + 4\mathbb{E}(\Delta_r^3)\mathbb{E}(B_r)\right) \\
&= \sum_r^{T_{<t}} (\delta^2 + 6\delta r),
\end{aligned}$$

since $\mathbb{E}(\Delta_r^3) = 0$ and $\mathbb{E}(\Delta_r^4) = \delta^2$. Finally,

$$\begin{aligned}
\mathbb{E}(B_t^4) &= \frac{t}{\delta}\delta^2 + 6\delta \sum_r^{T_{<t}} r = t\delta + 6\delta\left(\delta \sum_r^{{}^*\mathbb{N}_{<\frac{t}{\delta}}} r\right) \\
&= t\delta + 6\delta^2\left(\frac{\frac{t}{\delta}\left(\frac{t}{\delta} - 1\right)}{2}\right) \\
&= t\delta + 3t^2 - 3t\delta \\
&\leq 3t^2.
\end{aligned}$$

□

We can now prove the central result of our construction of the Wiener measure.

Proposition 3.3.9. *The internal stochastic process $B_{\lfloor t \rfloor_\delta}$ is S-continuous for P_L-almost all $\omega \in \Omega = \{-1, 1\}^T$. That is there is a set Ω' of full Loeb measure ($P_L(\Omega') = 1$) such that $B_{\lfloor t \rfloor_\delta}(\omega) \simeq B_{\lfloor t' \rfloor_\delta}(\omega)$ whenever for every $t \in [0, 1]$, $\omega \in \Omega'$, and $t' \in {}^*[0, 1]$ such that $t \simeq t'$. Moreover, $b_t(\omega)$ is continuous with respect to t for each $\omega \in \Omega'$.*

Recall that $\lfloor x \rfloor_\delta$ stand for the greatest multiple of δ smaller than x. The proof is a mix of the one in (Albeverio et al. 1986) and the original proof in (Anderson 1976).

Proof. We begin by showing that the set of paths Ω'' such that B_t is not S-continuous has measure zero. First we need some workable expression for Ω''. To this end, observe the following

$$\begin{aligned}
\Omega'' &= \left\{\omega \in \Omega \middle| \exists t \in [0, 1] \exists t' \in {}^*[0, 1](t \simeq t' \wedge B_{\lfloor t \rfloor_\delta}(\omega) \not\simeq B_{\lfloor t' \rfloor_\delta}(\omega))\right\} \\
&= \left\{\omega \in \Omega \middle| \exists m \in \mathbb{N} \exists t \in [0, 1] \exists t' \in {}^*[0, 1]\left(t \simeq t' \wedge \left|(B_{\lfloor t \rfloor_\delta} - B_{\lfloor t' \rfloor_\delta})(\omega)\right| \geq \frac{1}{m}\right)\right\} \\
&= \bigcup_m^{\mathbb{N}} \left\{\omega \in \Omega \middle| \forall n \in \mathbb{N} \exists i \in n \exists t \in \left[\frac{i}{n}, \frac{i+1}{n}\right] \exists t' \in {}^*\left[\frac{i}{n}, \frac{i+1}{n}\right]\left(\left|(B_{\lfloor t \rfloor_\delta} - B_{\lfloor t' \rfloor_\delta})(\omega)\right| \geq \frac{1}{m}\right)\right\} \\
&= \bigcup_m^{\mathbb{N}} \bigcap_n^{\mathbb{N}} \left\{\omega \in \Omega \middle| \exists i \in n \exists t \in \left[\frac{i}{n}, \frac{i+1}{n}\right] \exists t' \in {}^*\left[\frac{i}{n}, \frac{i+1}{n}\right]\left(\left|(B_{\lfloor t \rfloor_\delta} - B_{\lfloor t' \rfloor_\delta})(\omega)\right| \geq \frac{1}{m}\right)\right\} \\
&\subseteq \bigcup_m^{\mathbb{N}} \bigcap_n^{\mathbb{N}} \left\{\omega \in \Omega \middle| \exists i \in n\left(\left(\underset{t \in {}^*\left[\frac{i}{n}, \frac{i+1}{n}\right]}{\text{sup - inf}}\right) B_t(\omega) \geq \frac{1}{m}\right)\right\}
\end{aligned}$$

Let

$$\Omega_{m,n} = \left\{ \omega \in \Omega \middle| \exists i \in n \left(\left(\operatorname*{sup - inf}_{t \in {}^*\left[\frac{i}{n}, \frac{i+1}{n}\right]} \right) B_t(\omega) \geq \frac{1}{m} \right) \right\}.$$

Now,

$$\begin{aligned}
P(\Omega_{m,n}) &= P\left(\left\{ \omega \in \Omega \middle| \exists i \in n \left(\left(\operatorname*{sup - inf}_{t \in {}^*\left[\frac{i}{n}, \frac{i+1}{n}\right]} \right) B_t(\omega) \geq \frac{1}{m} \right) \right\} \right) \\
&= P\left(\bigcup_i^n \left\{ \omega \in \Omega \middle| \left(\left(\operatorname*{sup - inf}_{t \in {}^*\left[\frac{i}{n}, \frac{i+1}{n}\right]} \right) B_t(\omega) \geq \frac{1}{m} \right) \right\} \right) \\
&\leq \sum_i^n P\left(\left\{ \omega \in \Omega \middle| \left(\operatorname*{sup - inf}_{t \in {}^*\left[\frac{i}{n}, \frac{i+1}{n}\right]} \right) B_t(\omega) \geq \frac{1}{m} \right\} \right),
\end{aligned}$$

since P is subadditive. Moreover,

$$P(\Omega_{m,n}) \leq nP\left(\left\{ \omega \in \Omega \middle| \left(\operatorname*{sup - inf}_{t \in {}^*\left[\frac{0}{n}, \frac{1}{n}\right]} \right) B_t(\omega) \geq \frac{1}{m} \right\} \right)$$

since

$$P\left(\left\{ \omega \in \Omega \middle| \left(\operatorname*{sup - inf}_{t \in {}^*\left[\frac{0}{n}, \frac{1}{n}\right]} \right) B_t(\omega) \geq \frac{1}{m} \right\} \right) = P\left(\left\{ \omega \in \Omega \middle| \left(\operatorname*{sup - inf}_{t \in {}^*\left[\frac{i}{n}, \frac{i+1}{n}\right]} \right) B_t(\omega) \geq \frac{1}{m} \right\} \right),$$

for each $i \in n$ by translation symmetry. Furthermore,

$$\begin{aligned}
P(\Omega_{m,n}) &\leq nP\left(\left\{ \omega \in \Omega \middle| \left(\left| \sup_{t \in {}^*\left[\frac{0}{n}, \frac{1}{n}\right]} B_{\lfloor t \rfloor_\delta}(\omega) \right| + \left| \inf_{t \in {}^*\left[\frac{0}{n}, \frac{1}{n}\right]} B_{\lfloor t \rfloor_\delta}(\omega) \right| \geq \frac{1}{m} \right) \right\} \right) \\
&\leq nP\left(\left\{ \omega \in \Omega \middle| \left(2 \sup_{t \in {}^*\left[\frac{0}{n}, \frac{1}{n}\right]} \left| B_{\lfloor t \rfloor_\delta}(\omega) \right| \geq \frac{1}{m} \right) \right\} \right),
\end{aligned}$$

since,

$$\sup A - \inf A \leq |\sup A - \inf A| \leq |\sup A| + |\inf A| \leq 2 \sup_{a \in A} |a|,$$

for any internal set $A \subseteq {}^*\mathbb{R}$, by transfer. Now,

$$\begin{aligned}
P(\Omega_{m,n}) &\leq nP\left(\left\{ \omega \in \Omega \middle| \left(2 \max_{t \in T_{\leq \lfloor \frac{1}{n} \rfloor_\delta}} |B_t(\omega)| \geq \frac{1}{m} \right) \right\} \right) \\
&\leq nP\left(\left\{ \omega \in \Omega \middle| \left(\max_{t \in T_{\leq \lfloor \frac{1}{n} \rfloor_\delta}} B_t(\omega) \geq \frac{1}{2m} \right) \right\} \right) \\
&\quad + nP\left(\left\{ \omega \in \Omega \middle| \left(\min_{t \in T_{\leq \lfloor \frac{1}{n} \rfloor_\delta}} B_t(\omega) \leq -\frac{1}{2m} \right) \right\} \right) \\
&\leq n2P\left(\left\{ \omega \in \Omega \middle| \left(\max_{t \in T_{\leq \lfloor \frac{1}{n} \rfloor_\delta}} B_t(\omega) \geq \frac{1}{2m} \right) \right\} \right),
\end{aligned}$$

since

$$P\left(\left\{\omega \in \Omega \middle| \left(\max_{t \in T_{\leq \lfloor \frac{1}{n} \rfloor_\delta}} B_t(\omega) \geq \frac{1}{2m}\right)\right\}\right) = P\left(\left\{\omega \in \Omega \middle| \left(\min_{t \in T_{\leq \lfloor \frac{1}{n} \rfloor_\delta}} B_t(\omega) \leq -\frac{1}{2m}\right)\right\}\right),$$

by symmetry. Moreover,

$$P(\Omega_{m,n}) \leq 4nP\left(\left\{\omega \in \Omega \middle| \left(B_{\lfloor \frac{1}{n} \rfloor_\delta} \geq \frac{1}{2m}\right)\right\}\right)$$

since for each $\omega \in \Omega$ such that the maximum $\max\{|B_t(\omega)| \mid t \in T_{\leq \lfloor \frac{1}{n} \rfloor_\delta}\}$ at a $t < \lfloor \frac{1}{n} \rfloor_\delta$ we have a $\omega' \in \Omega$ such that $B_{\lfloor \frac{1}{n} \rfloor_\delta}(\omega') \geq \frac{1}{2m}$. Namely, let $\omega'(t') = \omega(t')$ for $t' \in T$ up to $t' = t$ and $\omega'(t') = -\omega(t')$ for $t' \in T_{\leq \lfloor \frac{1}{n} \rfloor_\delta}$ such that $t' > t$. Furthermore,

$$P(\Omega_{m,n}) \leq 4n(2m)^4 \mathbb{E}\left(B^4_{\lfloor \frac{1}{n} \rfloor_\delta}\right),$$

by Markov's Inequality (3.3.7), Finally, applying Lemma 3.3.8 we find that,

$$P(\Omega_{m,n}) \leq 4n(2m^4)3\left\lfloor \frac{1}{n} \right\rfloor_\delta^2 \simeq \frac{12(2m^4)}{n}.$$

Using this bound we can now prove that the P_L measure of the set Ω'' has measure zero. Observe that

$$P_L(\Omega'') \leq P_L\left(\bigcup_m^{\mathbb{N}} \bigcap_n^{\mathbb{N}} \Omega_{m,n}\right) = \lim_{m \to \infty} \lim_{n \to \infty} P_L(\Omega_{m,n}),$$

since $\Omega_{m+1,n} \supseteq \Omega_{m+1,n}$ and $\Omega_{m,n} \subseteq \Omega_{m,n+1}$ for each $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Finally,

$$P_L(\Omega'') \leq \lim_{m \to \infty} \lim_{n \to \infty} \frac{12(2m)^4}{n} = 0.$$

It follows that $\Omega' = \Omega \setminus \Omega''$ has full measure. That is the set on which $B_{\lfloor t \rfloor_\delta}$ is S-continuous has full P_L-measure. Moreover, on this set b_t is continuous with respect to t, by Proposition 3.3.4. $\square$

Proposition 3.3.10. *Suppose $\mathcal{B}$ is the Borel algebra of $C[0,1]$. Then the function $W : \mathcal{B} \to \mathbb{R}$ defined by*

$$W(A) := P_L(\{\omega \in \Omega \mid B(\omega) \in st^{-1}[A]\}),$$

for each $A \in \mathcal{B}$ is the Wiener measure. Here we interpret $B(\omega)$ as a function on $t \in T$,

$$B(\omega) : T \to {}^*\mathbb{R}; t \mapsto B_t(\omega).$$

Proof. First note that $st^{-1}[C[0,1]]$ is the set of nearstandard functions in ${}^*C[0,1]$ and that these are precisely the functions the S-continuous functions on ${}^*[0,1]$ (i.e. $st^{-1}[C[0,1]] = SC[0,1]$), by Proposition 3.3.5. It follows that,

$$W(C[0,1]) = P_L(\{\omega \in \Omega \mid B(\omega) \in \mathrm{st}^{-1}[C[0,1]]\}) = P_L(\Omega') = 1,$$

by Proposition 3.3.9.

Next we prove that each Borel set is measurable. We prove this first for the open balls in $C[0,1]$. Let $f \in C[0,1]$ and $\varepsilon \in \mathbb{R}_{>0}$. Now,

$$\begin{aligned}
W(B(f,\varepsilon)) &= P_L(\{\omega \in \Omega \mid B(\omega) \in \mathrm{st}^{-1}[B(f,\varepsilon)]\}) \\
&= P_L\left(\left\{\omega \in \Omega \middle| B(\omega) \in SC[0,1] \cap \bigcup_n^{\mathbb{N}} {}^*B\left(f, \varepsilon - \frac{1}{n}\right)\right\}\right) \\
&= P_L\left(\Omega' \cap \left\{\omega \in \Omega \middle| \bigcup_n^{\mathbb{N}} {}^*B\left(f, \varepsilon - \frac{1}{n}\right)\right\}\right) \\
&= P_L\left(\Omega' \cap \bigcup_n^{\mathbb{N}} \left\{\omega \in \Omega \middle| B(\omega) \in {}^*B\left(f, \varepsilon - \frac{1}{n}\right)\right\}\right).
\end{aligned}$$

It is not hard to see that the measurable sets form a σ-algebra and thus that all Borel sets are measurable.

To prove that W has the remaining two defining properties of the Wiener measure we make use of the fact that

$$\{\omega \in \Omega' \mid b(\omega) \in A\} = \{\omega \in \Omega \mid B(\omega) \in \mathrm{st}^{-1}[A]\}.$$

Now, let $f_t : C[0,1] \to [0,1]; h \mapsto h(t)$. Then

$$\begin{aligned}
W((f_t - f_s)^{-1}[(a,b]]) &= P_L(\{\omega \in \Omega \mid B(\omega) \in \mathrm{st}^{-1}[(f_t - f_s)^{-1}[(a,b]]\}) \\
&= P_L(\{\omega \in \Omega \mid b(\omega) \in (f_t - f_s)^{-1}[(a,b]]\}) \\
&= P_L(\{\omega \in \Omega' \mid \omega \in (b_t - b_s)^{-1}[(a,b]]\}) \\
&= P_L(\Omega' \cap (b_t - b_s)^{-1}[(a,b]]\}) \\
&= P_L((b_t - b_s)^{-1}[(a,b]]\}) \\
&= \frac{1}{\sqrt{2\pi(t-s)}} \int_a^b e^{-\frac{\xi^2}{2(t-s)}} d\xi
\end{aligned}$$

and

$$\begin{aligned}
&W\left(\bigcap_i^n (f_{t_i} - f_{s_i})^{-1}[(a_i,b_i]]\right) = P_L\left(\left\{\omega \in \Omega \middle| B(\omega) \in \mathrm{st}^{-1}\left[\bigcap_i^n (f_{t_i} - f_{s_i})^{-1}[(a_i,b_i]]\right]\right\}\right) \\
&= P_L\left(\left\{\omega \in \Omega \middle| b(\omega) \in \bigcap_i^n (f_{t_i} - f_{s_i})^{-1}[(a_i,b_i]]\right\}\right) \\
&= P_L\left(\left\{\omega \in \Omega' \middle| \omega \in \bigcap_i^n (b_{t_i} - b_{s_i})^{-1}[(a_i,b_i]]\right\}\right) \\
&= P_L\left(\Omega' \cap \bigcap_i^n (b_{t_i} - b_{s_i})^{-1}[(a_i,b_i]]\right) \\
&= P_L\left(\bigcap_i^n (b_{t_i} - b_{s_i})^{-1}[(a_i,b_i]]\right) = \prod_i^n P_L\left((b_{t_i} - b_{s_i})^{-1}[(a_i,b_i]]\right) \\
&= \prod_i^n P_L\left(\Omega' \cap (b_{t_i} - b_{s_i})^{-1}[(a_i,b_i]]\right) = \prod_i^n P_L\left((f_{t_i} - f_{s_i})^{-1}[(a_i,b_i]]\right).
\end{aligned}$$

□

Bibliography

Albeverio, Sergio et al. (1986). *Nonstandard Methods in Stochastic Analysis and Mathematical Physics*. Orlando, FL: Academic Press. ISBN: 0-12-048860-4.

Anderson, Robert M. (1976). "A non-standard representation for Brownian motion and Itô integration". In: *Israel Journal of Mathematics* 25.1-2, pp. 15–46. DOI: 10.1007/BF02756559. URL: https://doi.org/10.1007/BF02756559.

Friedman, Avner (2010). *Foundations of Modern Analysis*. Dover Books on Mathematics. Dover Publications. ISBN: 0-486-64062-0.

Henson, C. Ward (1979). "Unbounded Loeb measures". In: *Proceedings of the American Mathematical Society* 74.1, pp. 143–150. DOI: 10.1090/S0002-9939-1979-0521888-9. URL: https://doi.org/10.1090/S0002-9939-1979-0521888-9.

Loeb, Peter A. (1975). "Conversion from nonstandard to standard measure spaces and applications in probability theory". In: *Transactions of the American Mathematical Society* 211, pp. 113–122. DOI: 10.1090/S0002-9947-1975-0390154-8. URL: https://doi.org/10.1090/S0002-9947-1975-0390154-8.

— (1984). "A functional approach to nonstandard measure theory". In: *Contemporary Mathematics* 26, pp. 251–261.

Loeb, Peter A. and Robin Hurd (1985). *An Introduction to Nonstandard Analysis*. Orlando, FL: Academic Press. ISBN: 978-0124555600.

Loeb, Peter A. and Horst Osswald (1997). "Nonstandard integration theory in topological vector lattices". In: *Monatshefte für Mathematik* 124.1, pp. 53–82. DOI: 10.1007/BF01320737. URL: https://link.springer.com/article/10.1007/BF01320737.

Loeb, Peter A., Manfred P. H. Wolff, et al. (2000). *Nonstandard Analysis for the Working Mathematician*. Springer. ISBN: 978-94-017-7626-3.

Stroyan, K. D. and J. M. Bayod (1986). *Foundations of Infinitesimal Stochastic Analysis*. Vol. 119. Studies in Logic and the Foundations of Mathematics. Amsterdam: North-Holland Publishing Co. ISBN: 9780444879271.

Stroyan, K. D. and W. A. J. Luxemburg (1976). *Introduction to the Theory of Infinitesimals*. Vol. 72. Pure and Applied Mathematics. Academic Press. ISBN: 978-0126741506.