MASTERARBEIT | MASTER'S THESIS

Titel | Title

Stability in Continuous Logic

verfasst von | submitted by

Cezar Port

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 821
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Mathematik
Betreut von | Supervisor:	Univ.-Prof. Matthias Aschenbrenner PhD

Abstract

This thesis presents the basics of continuous model theory, gives an introduction to stability in continuous logic and ends with proving that the theory of Atomless Probability Algebras (APA) is $\aleph_0$-stable. For the basics we systematically follow an approach via type-spaces. Many continuous counterparts of the classical theorems from discrete model theory are proven, and we cover topics such as monster model, definability, algebraicity, conservative extensions of a theory, back-and-forth systems and quantifier elimination. We prove a lemma showing that formulas are approximable by unnested formulas, which in turn allows for more elementary proofs of some fundamental results than those given in [BBHU08].

We present three different proofs of the Fundamental Theorem of Stability which is the equivalence between the stability of a formula φ and the definablity of every φ-type. The first proof is original and elementary and only uses the definition of types. The second proof is a topological proof which follows the treatment in [BY14]. The third proof, communicated to me by Itaï Ben Yaacov, is an elementary approach that makes use of the properties of nets in a topological space.

In the final part of the thesis we provide a treatment with full proofs of the basic properties of the continuous theory of Probability Algebras Pr as in [BH23], with small variations. We provide an elementary proof of the fact that every probability algebra arises as the corresponding probability algebra of a probability space. We prove that every completion of Pr has QE, is separably categorical, and that every model of Pr is $\aleph_0$-saturated. This is very helpful in giving a complete proof of the $\aleph_0$-stability of APA.

Abstract [DE]

Diese Arbeit präsentiert die Grundlagen der stetigen Modelltheorie, gibt eine Einführung in den Stabilitätsbegriff in der stetigen Logik, und beweist abschließend, dass die Theorie der atomfreien Wahrscheinlichkeitsalgebren (APA) $\aleph_0$-stabil ist. Für die Grundlagen verfolgen wir systematisch einen Ansatz über Typenräume. Viele stetige Gegenstücke zu den klassischen Sätzen der diskreten Modelltheorie werden bewiesen. Wir behandeln Themen wie Monstermodell, Definierbarkeit, Algebraizität, konservative Erweiterungen einer Theorie, Back-and-Forth-Systeme, und Quantorenelimination. Wir beweisen ein Lemma, das zeigt, dass Formeln durch unverschachtelte Formeln approximierbar sind. Dies ermöglicht wiederum elementarere Beweise für einige fundamentale Ergebnisse als die in [BBHU08] gegebenen.

Wir präsentieren drei verschiedene Beweise für den Fundamentalsatz der Stabilitätstheorie, welcher die Äquivalenz zwischen der Stabilität einer Formel φ und der Definierbarkeit jedes φ-Typs besagt. Der erste Beweis ist neu und elementar und verwendet lediglich die Definition des Typbegriffs. Der zweite Beweis ist topologisch, und folgt der Darstellung in [BY14]. Der dritte Beweis, der mir von Itaï Ben Yaacov übermittelt wurde, ist ein elementarer Ansatz, der auf Eigenschaften von Netzen in einem topologischen Raum basiert.

Im letzten Teil der Arbeit geben wir eine Darstellung, mit vollständigen Beweisen, der grundlegenden Eigenschaften der stetigen Theorie der Wahrscheinlichkeitsalgebren Pr wie in [BH23], mit kleinen Variationen. Wir liefern einen elementaren Beweis dafür, dass jede Wahrscheinlichkeitsalgebra als die einem Wahrscheinlichkeitsraum assoziierte Wahrscheinlichkeitsalgebra entsteht. Wir beweisen, dass jede Vervollständigung von Pr QE erlaubt, separabel kategorisch ist, und dass jedes Modell von Pr $\aleph_0$-saturiert ist. Dies ist sehr hilfreich für den abschliessenden, vollständigen Beweis der $\aleph_0$-Stabilität von APA.

Contents

Introduction

This thesis provides a concise introduction to continuous model theory, beginning with its foundational concepts. It then presents a detailed treatment of the Fundamental Theorem of Stability, including a new and elementary proof demonstrating that the stability implies the definability of types. The thesis concludes with a study of the model theory of probability algebras, where many of the concepts and definitions introduced in earlier chapters are applied and further explored.

A first systematic definition and study of continuous logic was given by Chang and Keisler in [CK66]. However, their framework was too general in order to develop a continuous model theory which directly parallels the results of discrete model theory, where formulas take truth values in $\{0, 1\}$. In the logic introduced in [CK66], the truth value can be any element from a compact Hausdorff space. Later, in the late 2000's, in the article [BBHU08], continuous logic was developed with the set of truth values confined to the unit interval $[0, 1]$. This modification led to a more robust and complete theory, and the article [BBHU08] has since become the standard reference for studying continuous model theory. In this framework, the equality symbol is replaced by a binary relation symbol d, interpreted as a metric on the structures. For this reason, this logic is also called the model theory of metric structures.

This thesis is organized into four chapters. The first chapter, titled "Continuous Logic", uses extensively the articles [BYU10] and [BBHU08] in order to present the basics of continuous model theory. It begins with the section Syntax, which introduces the syntactic framework of continuous logic. The non-quantifier connectives are defined as all the continuous maps $[0, 1]^n \to [0, 1]$, while the quantifier connectives are inf and sup, which take the role of $\exists$ and $\forall$. Then we highlight two countable sets of connectives, one of which is fixed for use in the subsequent sections. We give the definition of a continuous logic language $\mathcal{L}$ as a triple $(\mathcal{L}^r, \mathcal{L}^f, (\Delta_s)_{s\in\mathcal{L}^r\cup\mathcal{L}^f})$ of relation symbols, function symbols, and for each symbol s an associated modulus of continuity Δ_s. The section concludes with the definition of continuous logic formulas and introduces the notations related to formulas and multivariables.

The section "Semantics" begins with the definition of $\mathcal{L}$-pre-structures $\boldsymbol{M}$, in which the interpretation of $d^{\boldsymbol{M}}$ is only a pseudo-metric. A countable system of connectives is fixed, and we prove that every definition is independent of the system of connectives chosen. Then we develop all the continuous model theory equivalents of the usual definitions in discrete model theory, such as a theory, elementary maps, and also introduce the $\mathcal{L}$-conditions, which serve the role of sentences. An $\mathcal{L}$-structure is defined as an $\mathcal{L}$-pre-structures in which $d^{\boldsymbol{M}}$ is a complete metric. This section ends with an important result showing that every $\mathcal{L}$-pre-structure can be embeded elementarily into an $\mathcal{L}$-structure.

The third section "Basic Theorems" presents full proofs of the continuous counterparts of all fundamental theorems of discrete model theory. We begin with the Tarski-Vaught Criterion and we introduce the size of an $\mathcal{L}$-structure which (unlike in classical logic where size equals cardinality) equals the cardinality of the smallest dense subset of metric space $(\boldsymbol{M}, d^{\boldsymbol{M}})$. The section proceeds with proving the Elementary Chain Theorem, Downward Löwenheim-Skolem Theorem, Łoś's Theorem, Compactness Theorem and concludes with the Upward Löwenheim-Skolem Theorem. The section ends by showing that every $\mathcal{L}$-formula can be approximated by unnested $\mathcal{L}$-formulas, a fact which is used in many results in the second chapter.

The final section "Type Spaces" is the most important as it sets the stage for later work. Here, we introduce the notion of types and type spaces $\mathrm{S}_x(T), \mathrm{S}_x(M)$ for a theory T, respectively for an $\mathcal{L}$-structure $\boldsymbol{M}$. We prove the essential theorem that $\mathrm{S}_x(T)$ can be equipped with a natural topology, called the logic topology, under which it is a compact hausdorff space. For any infinite cardinal κ, we define κ-saturated models, strongly κ-homogeneous models, and prove the existence of monster models. Then we show that $\mathrm{S}_x(T)$ can also be viewed as a complete metric space, whose topology is finer than the logic topology. One central idea of this section, is that we also show how to interpret an $\mathcal{L}$-formula φ as a continuous map $\dot{\varphi}$ from $\mathrm{S}_x(T) \to [0,1]$. We prove results on the behavior of this correspondence with respect to the continuous connectives. Here is also included a small section, named the "The logic mechanism on type spaces", which further develops this perspective. There we show how to also interpret $\mathcal{L}$-terms t as maps $\dot{t}$ from types spaces to type spaces, which in turn helps us understand better how $\dot{\varphi}$ breaks down as composition of maps $\dot{\psi}$ and $\dot{t}$, where ψ are $\mathcal{L}$-formulas and $\dot{t}$ are $\mathcal{L}$-terms.

The second chapter titled "Model Theory" builds on the section "Type Spaces" by presenting the fundamentals of continuous model theory. A central theme of this chapter is the consistent effort to define every concept through the lens of type spaces. For instance, in the section "Definability", we introduce the definable predicates $P : M_x \to [0,1]$ not as a cauchy limit of formulas as in [BBHU08], but as maps for which there exists a continuous function $\Phi : \mathrm{S}_x(T) \to [0,1]$ such that $P = \Phi \circ \iota$ where $\iota : M_x \to \mathrm{S}_x(T)$ is the natural map defined by $a \to \mathrm{tp}(a)$. We then study in detail how definable predicates interact with the continuous logic connectives, clarifying for the reader that these should be viewed precisely as formulas or as relation symbols added to the language. Additionally, we also present a simple topological proof of the continuous variant of Svenonius' Theorem.

Next, we introduce the definable sets, with a distinctive approach: first, we define definable sets of the type space $\mathrm{S}_x(T)$ which are subsets of the type space $\mathrm{S}_x(T)$, and subsequently define definable sets of the structure $\boldsymbol{M} \models T$ which uniquely correspond to the definable sets of the type space $\mathrm{S}_x(T)$. Then, we present the definable functions. Unfortunately, here we were not able to maintain a type-space based approach and we present the material following the treatment in [BBHU08]. Nonetheless, we establish several intuitive connections between definable sets and definable functions. This section ends with an elementary proof of the fact that if an $\mathcal{L}$-theory T is a conservative extension of another theory $\mathcal{L}_0$-theory T_0 and every non-logical symbol is defined in T over $\mathcal{L}_0$ then T is an extension by definitions of T_0.

In the section "Algebraicity", we introduce both definable and algebraic elements, prove the basic properties associated with them, and give an elementary proof of the fact that $\mathrm{acl}(\emptyset) = \mathrm{acl}(\mathrm{acl}(\emptyset))$. While in discrete model theory this result is almost trivial, in continuous model theory it required significantly more care. In [WH07] the authors use the monster model in order to prove $\mathrm{acl}(\emptyset) = \mathrm{acl}(\mathrm{acl}(\emptyset))$, here we give an elementary approach using only formulas.

The final section of this chapter "Topics in Model Theory", gathers several essential results. It begins with a topological lemma, which is at the basis of Chapter 3. Then we show that if there is a back-and-forth system between two structures, then every map from the back-and-forth system is elementary. This section concludes with a quantifier elimination criteria. Many results from this section are important for Chapter 4, dealing with probability algebras.

In the third chapter, titled "The Fundamental Theorem of Stability", we present the basics of stability theory and present two proofs of the fact that stability of a formula φ is equivalent to all the φ-types being definable (here-after called FTS). The first proof of FTS, presented in section 'Basics of Stability Theory'", is an original and elementary proof which only uses the definition of types. Rather than uniformly approximating the definable predicate $d_\varphi p$ with definable φ^{op}-predicates (as done in [BYU10]), the strategy is to explicitly construct and define $d_\varphi p$, and then prove its continuity. Specifically, I define the function

$$d_\varphi p : \mathrm{S}_{\varphi^{op}}(M) \to [0, 1]$$

and show that it is continuous. This suffices to conclude that $d_\varphi p$ is a definable φ^{op}-predicate.

The second proof presented in the section "Topological Proof of the Fundamental Theorem of Stability" follows the treatment in [BY14], and uses non-trivial topological results, based on an article by Grothendieck [Gro52]. The section "Convergence in Banach Spaces" presents with complete proofs the required topological results from [Gro52], and the main theorem is about the equivalence between a set being weakly relatively compact and a double limit property being satisfied.

Finally, the concluding section "Additional Results" of the chapter collects several additional results that follow as consequences of a formula being stable. We also introduce an equivalent formulation of stability, referred to as *approximately stable*, and show a third proof of FTS, communicated to me by Itaï Ben Yaacov, which is an elementary approach that makes use of the properties of nets in topological spaces.

In the final Chapter of this thesis "Application to Probability", we explore the fundamentals of the continuous theory of probability algebras, following the framework developed in [BH23]. The first section "Probability Algebras" introduces the definitions and basic properties of probability algebras and presents an elementary constructive proof that every probability algebra is isomorphic to a probability algebra corresponding to a probability space.

In the section "Model Theory of Probability Algebras", we present the theory Pr which axiomatizes probability algebras and prove that every completion of Pr has quantifier elimination, is separably categorical, and that every model of Pr is $\aleph_0$-saturated.

In the section "Stability of APA" we introduce the theory of Atomless Probability Spaces. We study the structure of its types spaces, and conclude with a proof that APA is $\aleph_0$-stable.

Finally, in the section "Applications", we present three applications of the model theory of probability algebras to results in probability theory. The first is a simple result which has applications in the theory of approximate groups, the second is Maharam's Theorem, and the third establishes the equivalence between the stability of APA, the stability of ARV (the theory of atomless random variables), and the Ryll-Nardzewski Theorem characterizing sequences of $[0, 1]$-valued random variables as spreadable if and only if they are exchangeable.

Preliminaries

In this thesis m and n range over the set $\mathbb{N} = \{0, 1, 2, \dots\}$ of natural numbers.

We say a topological space X is *compact* if every open cover of X has a finite subcover. If X and Y are two topological spaces, we denote by $\mathrm{C}(X, Y)$ the set of all continuous functions $X \to Y$.

For a bounded pseudo-metric space (X, d_X), and a non-empty set I, the set X^I, unless otherwise stated, is equipped with the supremum pseudo-metric defined by $\max_i d_X(a_i, b_i)$ for $(a_i), (b_i) \in X^I$. By abuse of notation we will write d_X for the pseudo-metric on X^I.

Moreover, if (X, d_X) is a complete metric space and I is a non-empty set, then (X^I, d_X) is also a complete metric space. This is a consequence of the fact that if (a_n) is a cauchy sequence in X^I, then for each $i \in I$ the sequence $(a_n^i)_n$, formed by the i-th coordinates of (a_n), is a cauchy sequence in X.

Let Λ be a non-empty linearly ordered set and $(X_\lambda)_{\lambda \in \Lambda}$ be a family of pseudo-metric spaces such that for any $\lambda, \mu \in \Lambda$ with $\lambda \leq \mu$ we have $X_\lambda \subseteq X_\mu$ and $(X_\lambda, d_{X_\lambda})$ is a pseudo-metric subspace of (X_μ, d_{X_μ}). We can then define a pseudo-metric space X, called *the union of the increasing chain* $(X_\lambda)_\lambda$, by setting $X = \bigcup_\lambda X_\lambda$, and defining the distance between two elements $a, b \in X$ by $d_X(a, b) = d_{X_\lambda}(a, b)$ for any $\lambda \in \Lambda$ such that $a, b \in X_\lambda$. This is well-defined because the pseudo-metrics are compatible along the chain.

For a pseudo-metric space (X, d_X) with diamater ≤ 1, given a point $x \in X$ and a subset $A \subseteq X$, we define *the distance from x to A* by $d_X(x, A) = \inf\{d_X(x, a) : a \in A\}$. By convention, if $A = \emptyset$, we set $d_X(x, \emptyset) = 1$.

A subset $Y \subseteq X$ of a pseudo-metric space X is called an *ε-net* if every point in X is within distance at most ε from some point in Y, that is,

$$X = \bigcup_{a \in Y} B(a, \varepsilon),$$

where $B(a, \varepsilon) = \{b \in X \mid d(a, b) \leq \varepsilon\}$.

1 Continuous Logic

This chapter uses extensively the articles [BYU10] and [BBHU08] in order to present the basics of continuous logic. We assume basic knowledge of syntax and semantics of first-order logic.

1.1 Syntax

The language of continuous logic is formed of logical and non-logical symbols. The non-logical symbols are the function symbols and relation symbols, like in ordinary first-order logic, while the logical symbols are the connectives, which are specific to continuous logic, the quantifiers denoted by inf and sup, and a symbol d, taking a role similar to the equality symbol in classical logic.

As for connectives, we can take for every n all the continuous functions $[0,1]^n$ into $[0,1]$ (in classical logic these are functions from $\{0,1\}^n$ into $\{0,1\}$). Using this many connectives would give an uncountable number of sentences. Therefore, in a countable language to have a countable number of sentences, we need to restrict to a countable number of connectives, while still being able to express, in an approximate manner, all the statements that can be expressed using all possible connectives. Hence we define:

Definition 1.1. A *system of connectives* is a family $\mathcal{F} = \{F_n : n \in \mathbb{N}\}$ where $F_n \subseteq C([0,1]^n, [0,1])$. Elements of F_n are called *n-ary connectives*. The system $\mathcal{F}$ is said to be *closed* if:

(i) for any $m < n$ the projection $\pi_{n,m} : [0,1]^n \to [0,1]$ is in F_n;

(ii) if $f \in F_n$ and $g_0, \dots, g_{n-1} \in F_m$ then $f \circ (g_0, \dots, g_{n-1}) \in F_m$.

In the case $n = 0$ of Definition 1.1, f is a constant function $\{\emptyset\} \to [0,1]$ while $(g_0 \dots g_{n-1})$ is understood to be the function $[0,1]^m \to \{\emptyset\}$. Hence if the constant 0 belongs to F_0 for a closed system $\mathcal{F}$, then by Definition 1.1(ii), for every n the constant function $[0,1]^n \to \{0\}$ belongs to F_n.

For any system of connectives $\mathcal{F}$ we can consider the smallest closed system of connectives containing $\mathcal{F}$, denoted by $\overline{\mathcal{F}}$. Finally, we will say that a system of connectives $\mathcal{F}$ is *full* if for each $n \geq 1$ the set of n-ary connectives in $\overline{\mathcal{F}}$ is dense in the space $C([0,1]^n, [0,1])$ with the uniform topology.

Remark 1.2. *All connectives are uniformly continuous. The system of connectives $\mathcal{C} = \{\mathrm{C}([0,1]^n, [0,1]) : n \in \mathbb{N}\}$ is full.*

Let $\neg$ represent the unary connective $x \mapsto 1-x$, let $\frac{x}{2}$ represent the connective $x \mapsto \frac{x}{2}$, and $\dot{-}$ represent the binary connective $(x,y) \mapsto \max(0, x-y)$. Using $\{\neg, \frac{x}{2}, \dot{-}\}$ we can build connectives such as $x \dot{+} y = \neg(\neg x \dot{-} y)$, $\min(x,y) = x \dot{-} (x \dot{-} y)$, $|x-y| = (x \dot{-} y) \dot{+} (y \dot{-} x)$, and many others.

Theorem 1.3. *[BYU10, Corollary 1.5 and 1.6]*
1) For any dense $C \subseteq [0,1]$, the system of connectives

$$F_0 = C, F_1 = \emptyset, F_2 = \{\dot{-}\} \text{ and } F_n = \emptyset \quad \text{for } n > 2$$

is full.
2) $F_0 = \emptyset, F_1 = \{\neg, \frac{x}{2}\}, F_2 = \{\dot{-}\}$ and $F_n = \emptyset$ for $n > 2$ is a full system.

The proof of this theorem relies on the following consequence of the lattice version of the Stone-Weierstrass Theorem. We include it both for completeness and because it will be used multiple times in later results.

Proposition 1.4. *[BYU10, Proposition 1.4] Let X be a compact hausdorff space and let $\mathcal{S} \subseteq \mathrm{C}(X,[0,1])$. If $\mathcal{S}$ satisfies the following conditions:*

(i) it is closed under $\neg$ and $\dot{-}$,

(ii) it separates points in X, meaning that for every distinct $x, y \in X$, there exists $f \in \mathcal{S}$ such that $f(x) \neq f(y)$,

(iii) either $\mathcal{S}$ is closed under $\frac{x}{2}$, or the set of constant functions in $\mathcal{S}$ is dense in $[0,1]$,

then $\mathcal{S}$ is dense in $\mathrm{C}(X,[0,1])$.

Next we define a way to measure the uniform continuity of a function.

Definition 1.5. A *modulus of continuity* is a function $\delta : (0,\infty) \to (0,\infty)$. Let (X, d_X) and (Y, d_Y) be two pseudo-metric spaces. Then a map $f : X \to Y$ is said to be *uniformly continuous with respect to δ* if for all $\varepsilon > 0$ and $a, b \in X$ we have

$$d_X(a,b) < \delta(\varepsilon) \implies d_Y(f(a), f(b)) \leq \varepsilon.$$

We also say that f has modulus of continuity δ.

We now arrive at the main definition of this section:

Definition 1.6. A (continuous) language $\mathcal{L}$ is a triple $(\mathcal{L}^r, \mathcal{L}^f, (\Delta_s)_{s \in \mathcal{L}^r \cup \mathcal{L}^f})$ where $(\mathcal{L}^r, \mathcal{L}^f)$ is a (one-sorted) language in the usual sense of first-order logic, and Δ_s is a modulus of continuity for each non-logical symbol s. The size of $\mathcal{L}$ is defined to be

$$|\mathcal{L}| = \max(\aleph_0, |\mathcal{L}^r \cup \mathcal{L}^f|).$$

Let $\mathcal{L}$ be a language. We build $\mathcal{L}$-terms recursively, as in first-order logic, starting from the variables (for continuous logic we fix a countable set of quantifiable variables and a set of unquantifiable variables, exactly as in classical logic), and by induction $ft_1 \ldots t_n$ is an $\mathcal{L}$-term if f is an n-ary function symbol and each t_i is an $\mathcal{L}$-term.

Definition 1.7. For a system of connectives $\mathcal{F}$, we build the $\mathcal{L}$-formulas recursively:

(i) dt_1t_2 is an $\mathcal{L}$-formula for all $\mathcal{L}$-terms t_1, t_2;

(ii) $Rt_1 \dots t_n$ is an $\mathcal{L}$-formula for each n-ary relation symbol $R \in \mathcal{L}^r$ and $\mathcal{L}$-terms $t_1, \dots, t_n$;

(iii) for each γ an n-ary connective from $\mathcal{F}$ if $\varphi_1, \dots, \varphi_n$ are $\mathcal{L}$-formulas then $\gamma\varphi_1 \dots \varphi_n$ are $\mathcal{L}$-formulas;

(iv) for each $\mathcal{L}$-formula φ and quantifiable variable x, $\inf_x \varphi$ (denoting the string of symbols $\inf x\varphi$) and $\sup_x \varphi$ are $\mathcal{L}$-formulas.

In order to emphasize the system of connectives used for the construction of formulas we will use the notation of $\mathcal{L}_{\mathcal{F}}$-formulas. In [BBHU08, Definition 6.2] these are called $\mathcal{F}$-restricted $\mathcal{L}$-formulas. All the notions from classical logic, of sentences, bound variables, free variables, substitution of terms/formulas and quantifier-free formulas translate in a straightforward way to continuous languages.

Definition 1.8. A *multivariable* of $\mathcal{L}$ is a tuple $x = (x_i)_{i \in I}$, where I is an (possibly infinite) index set, of distinct variables of $\mathcal{L}$. For an $\mathcal{L}$-formula φ, we use the notation $\varphi(x)$ if the free variables of φ are among x, and we say φ is an $\mathcal{L}(x)$-formula. A multivariable y is called *quantifiable* if it is a collection of quantifiable variables.

Remark 1.9. *Let x be a multivariable, y be a quantifiable multivariable, and φ be an $\mathcal{L}(x)$-formula. Then by the expression* $\inf_y \varphi$ *we mean the formula* $\inf_z \varphi$*, where z is the finite multivariable containing all variables appearing in φ and in y. We then regard* $\inf_y \varphi$ *as a formula with free variables in $x \setminus y$.*

Definition 1.10. For a multivariable x, by abuse of notation, we denote by $\mathcal{L}(x)$ the set of all $\mathcal{L}(x)$-formulas. Moreover, the property of unique readability holds for terms and formulas as defined above.

1.2 Semantics

Let $\mathcal{L}$ be a language, for which we use the system of all connectives $\mathcal{C}$ (Remark 1.2).

Definition 1.11. An $\mathcal{L}$*-pre-structure* is a tuple $\boldsymbol{M} = (M, d^{\boldsymbol{M}}, (R^{\boldsymbol{M}})_{R \in \mathcal{L}^r}, (f^{\boldsymbol{M}})_{f \in \mathcal{L}^f})$ such that:

(i) $(M, d^{\boldsymbol{M}})$ is a nonempty pseudo-metric space with diameter ≤ 1;

(ii) for each relation symbol $R \in \mathcal{L}^r$ of arity n, $R^{\boldsymbol{M}}$ is a function $M^n \to [0, 1]$;

(iii) for each function symbol $f \in \mathcal{L}^f$ of arity n, $f^{\boldsymbol{M}}$ is a function $M^n \to M$;

(UC) for each $s \in \mathcal{L}^f \cup \mathcal{L}^r$, $s^{\boldsymbol{M}}$ is uniformly continuous with respect to the modulus of continuity Δ_s.

UC is an abbreviation for uniform continuity conditions. As stated in the preliminaries, for every non-empty set I, we will use $d^{\boldsymbol{M}}$ for the supremum pseudo-metric on M^I, hence the UC holds with respect to this pseudo-metric. The functions $d^{\boldsymbol{M}}, f^{\boldsymbol{M}}, R^{\boldsymbol{M}}$ are called the *interpretations* in $\boldsymbol{M}$ of the respective $\mathcal{L}$-symbols. If we want to expand $\boldsymbol{M}$ by names for elements of a parameter set $A \subseteq M$, then we define $\mathcal{L}_A$ to be $\mathcal{L}$ augmented by distinct new constant symbols for every element from A, while $\boldsymbol{M}_A$ will denote the $\mathcal{L}_A$-pre-structure with these new constant symbols interpreted in the obvious way.
For a multivariable $x = (x_i)_{i \in I}$, we define the x-set of $\boldsymbol{M}$ as $M_x := M^I$. The interpretation of an $(\mathcal{L}_\mathcal{C})_M$-term t, with no free-variables, in $\boldsymbol{M}$, is defined in a similar way as in classical logic and denoted by $t^{\boldsymbol{M}}$. The *truth value* of an $(\mathcal{L}_\mathcal{C})_M$-sentence σ, denoted by $\sigma^{\boldsymbol{M}}$, is defined recursively:

(i) $(dt_1t_2)^{\boldsymbol{M}} := d^{\boldsymbol{M}}(t_1^{\boldsymbol{M}}, t_2^{\boldsymbol{M}})$ for t_1, t_2 variable-free $(\mathcal{L}_\mathcal{C})_M$-terms;

(ii) $(Rt_1 \dots t_n)^{\boldsymbol{M}} := R^{\boldsymbol{M}}(t_1^{\boldsymbol{M}}, \dots t_n^{\boldsymbol{M}})$ for each $R \in \mathcal{L}^r$ and $t_1, \dots, t_n$ variable-free $(\mathcal{L}_\mathcal{C})_M$-terms;

(iii) for γ an n-ary connective and $(\mathcal{L}_\mathcal{C})_M$-sentence $\gamma\sigma_1 \dots \sigma_n$ we define

$$(\gamma\sigma_1 \dots \sigma_n)^{\boldsymbol{M}} := \gamma(\sigma_1^{\boldsymbol{M}}, \dots, \sigma_n^{\boldsymbol{M}});$$

(iv) for an $(\mathcal{L}_\mathcal{C})_M$-sentence $\sigma = \inf_y \varphi$, where $\varphi(y)$ is an $(\mathcal{L}_\mathcal{C})_M$-formula, y a quantifiable variable we define $\sigma^{\boldsymbol{M}} := \inf\{\varphi(a)^{\boldsymbol{M}} : a \in M_y\}$;

(v) for an $(\mathcal{L}_\mathcal{C})_M$-sentence $\sigma = \sup_y \varphi$, where $\varphi(y)$ is an $(\mathcal{L}_\mathcal{C})_M$-formula, y a quantifiable variable we define $\sigma^{\boldsymbol{M}} := \sup\{\varphi(a)^{\boldsymbol{M}} : a \in M_y\}$.

Therefore for an $\mathcal{L}_\mathcal{C}$-term t, if x contains all the free variables of t, we can define the function $t_x^{\boldsymbol{M}} : M_x \to M, a \mapsto t(a)^{\boldsymbol{M}}$, and for an $\mathcal{L}_\mathcal{C}$-formula φ, if x contains all the free variables of φ, we can define the function $\varphi_x^{\boldsymbol{M}} : M_x \to [0,1], a \mapsto \varphi(a)^{\boldsymbol{M}}$. It is clear that $\varphi_x^{\boldsymbol{M}}$ factors through $M_{(v_1,\dots,v_n)}$, where $v_1, \dots, v_n$ are the free variables of φ. Once a multivariable x is fixed, we will simply write $\varphi^{\boldsymbol{M}}$ and $t^{\boldsymbol{M}}$, omitting explicit reference to x.

Proposition 1.12. *For every $\mathcal{L}_\mathcal{C}$-formula $\varphi(x)$ the function $\varphi^{\boldsymbol{M}}$ has a modulus of continuity which depends only on φ and not on the $\mathcal{L}$-pre-structure $\boldsymbol{M}$.*

This proposition is a consequence of the fact that compositions of functions which have moduli of continuity, also have a modulus of continuity. The quantifiers $\inf, \sup$ even preserve the modulus of continuity.

Remark 1.13. *Given two $\mathcal{L}_\mathcal{C}$-formulas φ, ψ, if φ, ψ are $\mathcal{L}(x)$-formulas, we call the supremum of the expression*

$$\left|\varphi^{\boldsymbol{M}}(a) - \psi^{\boldsymbol{M}}(a)\right|,$$

where $a \in M_x$ and $\boldsymbol{M}$ ranges over all $\mathcal{L}$-pre-structures, the logical distance between φ and ψ. The logical distance is independent of x.

The logical distance is actually a pseudo-metric on the set of $\mathcal{L}_{\mathcal{C}}(x)$-formulas. Hence we arrive at the following definition:

Definition 1.14. Two $\mathcal{L}_{\mathcal{C}}$-formulas $\varphi(x), \psi(x)$ are said to be *logically equivalent* if the logical distance between them is zero in the pseudo-metric space of $\mathcal{L}_{\mathcal{C}}(x)$-formulas. Equivalently $\varphi^{\boldsymbol{M}} = \psi^{\boldsymbol{M}}$ for any $\mathcal{L}$-pre-structure $\boldsymbol{M}$.

The following theorem gives motivation for using a full system of connectives.

Theorem 1.15. *[BBHU08, Theorem 6.3] Let $\mathcal{F}$ be a full system of connectives. Then the set of $\mathcal{L}_{\mathcal{F}}(x)$-formulas is dense in the pseudo-metric space of $\mathcal{L}_{\mathcal{C}}(x)$-formulas.*

The proof uses the fact that the connectives are uniformly continuous functions, and performs an induction on the length of formulas. Therefore by using a countable full system of connectives for a countable language, we obtain in total a countable number of sentences, which mirrors the classical logic case, and we are able to express in an uniform approximate way any formula from $\mathcal{L}_{\mathcal{C}}$.

Convention 1.16. *From this point on we fix and use for each language $\mathcal{L}$ the full system of connectives $\mathcal{F} = \{\neg, \frac{x}{2}, \dot{-}\}$, unless otherwise stated. By an $\mathcal{L}$-formula we simply mean an $\mathcal{L}_{\mathcal{F}}$-formula. This does not restrict us in any way, as we will prove that every model theoretic notion is independent of the full system of connectives chosen. Here $\mathcal{F}$ is chosen out of convenience of being able to write expressions more succinctly. Also, we will use freely intuitive symbols for connectives which can be obtained by composition of connectives from $\mathcal{F}$, such as $\dot{+}, \min, \max, |x|, 0, 1$ and r where r is any dyadic number in $[0,1]$. As an example 0 is obtained by $\inf_x d(x,x)$.*

Definition 1.17. An *$\mathcal{L}$-condition* E is a formal expression $\varphi = 0$ where φ is an $\mathcal{L}$-formula. If φ is a sentence then E is called a *closed* condition. If $\varphi = \varphi(x)$ we denote the condition by $E(x)$ and call it an *$\mathcal{L}(x)$-condition.* For an $\mathcal{L}$-pre-structure $\boldsymbol{M}$ and $a \in M_x$ we define E *to be satisfied by* a in $\boldsymbol{M}$ if $\varphi^{\boldsymbol{M}}(a) = 0$. This is denoted by $\boldsymbol{M} \models E(a)$ or $a \models E$, if $\boldsymbol{M}$ is clear from the context. We will also say $E(a)$ holds in $\boldsymbol{M}$ or that $E(a)$ is true in $\boldsymbol{M}$.

There are a number of abbreviations used: $\varphi = \psi$ is an abbreviation for $(\varphi \dot{-} \psi) \dot{+} (\psi \dot{-} \varphi) = 0$ and $\varphi \leq \psi$ is an abbreviation for $\varphi \dot{-} \psi = 0$. Additionally, in order to draw a parallel with first-order logic, we will occasionally use the notations $\forall x$ and $\exists x$ to represent the continuous quantifiers $\sup_x$ and $\inf_x$, respectively.

Definition 1.18. In order to express arguments and ideas more succinctly, we will say that a formal expression E *is true in* $\boldsymbol{M}$ if $\boldsymbol{M}$ satisfies a certain set S of $\mathcal{L}$-conditions. Intuitively S is the set which expresses the "property" attached to E, and by abuse of notation we will denote denote S by E or "E" for clarity. For example, if $\varphi(x)$ is an $\mathcal{L}$-formula and $r \in [0,1]$, then the expression $\varphi \leq r$ can be "expressed" by the set of $\mathcal{L}$-conditions

$$\text{“}\varphi \leq r\text{”} = \{\varphi \leq d : r \leq d, d \in [0,1] \text{ a dyadic number}\}.$$

Therefore, for $a \in M_x$ we say $\varphi \leq r$ holds for a in $\boldsymbol{M}$, denoted by $\boldsymbol{M} \models \varphi(a) \leq r$, if $\boldsymbol{M} \models \varphi(a) \leq d$ for each dyadic $d \in [0,1]$ with $r \leq d$.

If $d^{\boldsymbol{M}}$ is a metric then the property $x = y$ is expressed by the $\mathcal{L}$-condition $d(x,y) = 0$. On the other hand the property $x \neq y$ cannot be expressed using a set of $\mathcal{L}$-conditions. Also, there is no way to express the property $\varphi < r$ using a set of $\mathcal{L}$-conditions.

Remark 1.19. *In continous logic, the numerical value assigned to true is* 0 *and to false is* 1*, as reflected in the definition of an $\mathcal{L}$-condition. This assignment is motivated by the following example: the condition $\varphi(a) = 0$ is true for all $a \in M_x$ is equivalent to $\forall x \varphi(x) = 0$ is true, while if we would assign $T = 1, F = 0$ then $\forall x \varphi(x) = 1$ being true would not imply that $\varphi(a) = 1$ is true for all $a \in M_x$ (this is because $\forall$ is the quantifier* sup*).*

Definition 1.20. Let Σ be a set of closed $\mathcal{L}$-conditions. Then for an $\mathcal{L}$-pre-structure $\boldsymbol{M}$ we say $\boldsymbol{M}$ *is a model of* Σ if $\boldsymbol{M}$ satisfies every condition from Σ, denoted by $\boldsymbol{M} \models \Sigma$. The *theory of* Σ, denoted by $\mathrm{Th}(\Sigma)$, is the set of all closed conditions which are true in any model of Σ. We call a set of closed $\mathcal{L}$-conditions T a theory if $T = \mathrm{Th}(T)$. The *theory of* $\boldsymbol{M}$, $\mathrm{Th}(\boldsymbol{M})$, is the set of all closed $\mathcal{L}$-conditions which are true in $\boldsymbol{M}$. We also denote $\varphi \in \mathrm{Th}(\Sigma)$ by $\Sigma \models \varphi$. Finally, we say that Σ is *satisfiable* if there is an $\mathcal{L}$-pre-structure which models Σ.

Definition 1.21. Let $\boldsymbol{M}, \boldsymbol{N}$ be two $\mathcal{L}$-pre-structures. We say $\boldsymbol{M}, \boldsymbol{N}$ are *elementarily equivalent*, and denote this by $\boldsymbol{M} \equiv \boldsymbol{N}$, if $\mathrm{Th}(\boldsymbol{M}) = \mathrm{Th}(\boldsymbol{N})$.

Hence $\boldsymbol{M} \equiv \boldsymbol{N}$ if and only if $\sigma^{\boldsymbol{M}} = \sigma^{\boldsymbol{N}}$ for any $\mathcal{L}$-sentence σ.

Remark 1.22. *Observe that since the set of $\mathcal{L}_{\mathcal{G}}$-sentences, for a full system of connectives $\mathcal{G}$, is dense in the set of $\mathcal{L}_{\mathcal{C}}$-sentences, with respect to the logical distance, the following are equivalent:*

(i) $\boldsymbol{M}$ and $\boldsymbol{N}$ satisfy the same $\mathcal{L}_{\mathcal{C}}$-conditions;

(ii) $\boldsymbol{M}$ and $\boldsymbol{N}$ satisfy the same $\mathcal{L}_{\mathcal{G}}$-conditions.

Therefore the condition $\boldsymbol{M} \equiv \boldsymbol{N}$ is independent of the full system of connectives chosen.

Definition 1.23. Let $\boldsymbol{M}, \boldsymbol{N}$ be two $\mathcal{L}$-pre-structures. A map $f : M \to N$ which preserves the interpretation of d, function and relation symbols on the structures is called an $\mathcal{L}$-*morphism*, and is denoted by $f : \boldsymbol{M} \to \boldsymbol{N}$.

Definition 1.24. Let $\boldsymbol{M}, \boldsymbol{N}$ be two $\mathcal{L}$-pre-structures, and let $A \subseteq M, B \subseteq N$ be subsets. A map $f : A \to B$ is called *elementary* if for every $\mathcal{L}$-condition $E(x)$ and all $a \in A_x$,

$$\boldsymbol{M} \models E(a) \Longleftrightarrow \boldsymbol{N} \models E(f(a)).$$

Since elementary maps $f : M \to N$ are in particular morphisms, we will abuse notation and also say $f : \boldsymbol{M} \to \boldsymbol{N}$ is an elementary map. We say that $\boldsymbol{M}$ is a *substructure* of $\boldsymbol{N}$, denoted by $\boldsymbol{M} \subseteq \boldsymbol{N}$, if $M \subseteq N$ and the natural inclusion $M \to N$ is a morphism

$\boldsymbol{M} \to \boldsymbol{N}$. If $\boldsymbol{M}$ is a substructure of $\boldsymbol{N}$ and the natural inclusion $M \to N$ is elementary, then we say $\boldsymbol{M}$ is an elementary substructure of $\boldsymbol{N}$, denoted by $\boldsymbol{M} \preceq \boldsymbol{N}$. Similarly as in the remark before, the notion of a map being elementary, or a substructure being elementary, is independent of the full system of connectives chosen.

Having a pseudo-metric, instead of a metric, does come at the expense that using conditions we are not able to force two elements in a pre-structure to be equal. Therefore we arrive at the main definition of this section:

Definition 1.25. An $\mathcal{L}$-pre-structure $\boldsymbol{M}$ for which $(M, d^{\boldsymbol{M}})$ is a complete metric space is called an *$\mathcal{L}$-structure.*

Remark 1.26. *The metric space $(M_x, d^{\boldsymbol{M}})$ is complete.*

In the following paragraphs we show how to attach in an unique way an $\mathcal{L}$-structure to an $\mathcal{L}$-pre-structure in two steps, which is a proof by example of Theorem 1.27. Let $\boldsymbol{M}$ be an $\mathcal{L}$-pre-structure with pseudo-metric $d^{\boldsymbol{M}}$. By identifying elements which are in zero distance from each other, an $\mathcal{L}$-pre-structure $\boldsymbol{M}_0$ is obtained with its distance being a metric. The underlying set of $\boldsymbol{M}_0$ is $M_0 = M/\{(a,b) \in M^2 : d^{\boldsymbol{M}}(a,b) = 0\}$, with the obvious induced metric $d^{\boldsymbol{M}_0}$. The functions and relations induced by $\boldsymbol{M}$ are well-defined on $\boldsymbol{M}_0$ as a consequence of the uniform continuity conditions. As an example, let an n-ary function $f^{\boldsymbol{M}} : M^n \to M$ and $a, b \in M^n$ with $d^{\boldsymbol{M}}(a,b) = 0$. Applying the UC, for all $\varepsilon > 0$, since $d^{\boldsymbol{M}}(a,b) < \Delta_f(\varepsilon)$ we get $d^{\boldsymbol{M}}(f^{\boldsymbol{M}}(a), f^{\boldsymbol{M}}(b)) \leq \varepsilon$. Hence $d^{\boldsymbol{M}}(f^{\boldsymbol{M}}(a), f^{\boldsymbol{M}}(b)) = 0$, which in turn implies that the interpretation of f is well defined on $\boldsymbol{M}_0$. Relations are well-defined by a similar argument. The UC are satisfied also for $\boldsymbol{M}_0$ with the induced metric, functions and relations, making $\boldsymbol{M}_0$ into an $\mathcal{L}$-pre-structure with its distance being a metric. Moreover, the natural projection $\pi : M \to M_0$ is an elementary map.

Next, we consider the completion $(\widehat{M}, d^{\widehat{\boldsymbol{M}}})$ of $(M_0, d^{\boldsymbol{M}_0})$. Again, the UC helps us to extend the functions and relations. Continuing the example from above, we have the function $f^{\boldsymbol{M}_0} : M_0^n \to M_0$ and we define $f^{\widehat{\boldsymbol{M}}} : \widehat{M}^n \to \widehat{M}$ for $a = (a^1, \ldots, a^n) \in \widehat{M}^n$ by picking any sequence $((a_j^1, \ldots, a_j^n))_{j\in\mathbb{N}} \subseteq M_0^n$ with $\lim_j a_j^i = a^i$ for each i and setting

$$f^{\widehat{\boldsymbol{M}}}(a) = \lim_j f^{\boldsymbol{M}_0}(a_j^1, \ldots, a_j^n).$$

Since $(a_j^1, \ldots, a_j^n)_{j\in\mathbb{N}}$ is a cauchy sequence in M_0^n, by UC, $(f^{\boldsymbol{M}_0}(a_j^1, \ldots, a_j^n))_j$ is a cauchy sequence in $\widehat{M}$, therefore the limit exists. Also, the UC implies that the limit is well-defined, i.e., independent of the chosen sequence $((a_j^1, \ldots, a_j^n))_{j\in\mathbb{N}}$. Obviously $f^{\widehat{\boldsymbol{M}}}$ extends $f^{\boldsymbol{M}_0}$. Relations are extended in a similar way.

The UC conditions are also satisfied by $\widehat{\boldsymbol{M}}$, and since $d^{\widehat{\boldsymbol{M}}}$ is a complete metric, it is actually an $\mathcal{L}$-structure. Moreover the natural injection $\iota : M_0 \to \widehat{M}$ is an elementary map $\boldsymbol{M}_0 \to \widehat{\boldsymbol{M}}$, proven by using Proposition 1.12. Summarizing, we obtain:

Theorem 1.27. *[BYU10, Proposition 2.10] Let $\boldsymbol{M}$ be an $\mathcal{L}$-pre-structure. Then with notations as above, there is a unique way to define an $\mathcal{L}$-structure, denoted by $\widehat{\boldsymbol{M}}$, with* *underlying metric space $(\widehat{M}, d^{\widehat{\boldsymbol{M}}})$ for which the natural mapping $\iota \circ \pi : M \to \widehat{M}$ is a morphism. Then $\iota \circ \pi$ is an elementary map, and any morphism $\boldsymbol{M} \to \boldsymbol{N}$, with $\boldsymbol{N}$ an $\mathcal{L}$-structure, factors uniquely through $\iota \circ \pi$. Moreover, $\iota(\pi(M)))$ is a dense subset of $\widehat{M}$.*

Uniqueness follows from the uniform continuity of the functions and relations.

Remark 1.28. *By Theorem 1.27, for a closed $\mathcal{L}$-condition E, $\boldsymbol{M} \models E$ if and only if $\widehat{\boldsymbol{M}} \models E$. Therefore the notions of* $\mathrm{Th}(\Sigma)$ *and $\Sigma \models \varphi$ from Definition 1.20 remain the same if we change in the definition the expression "$\mathcal{L}$-pre-structure" to "$\mathcal{L}$-structure".*

1.3 Basic Theorems

We start with a criterion which is the basis of many theorems and applications.

Theorem 1.29 (Tarski-Vaught Criterion)**.** *Let $\boldsymbol{M}$ be an $\mathcal{L}$-structure and $A \subseteq M$ closed and nonempty. The following are equivalent:*

(i) A underlies an elementary substructure of $\boldsymbol{M}$;

(ii) for each $\mathcal{L}$-formula $\varphi(x, y)$, with y a quantifiable variable, and every $a \in A_x$ we have

$$(\inf_y \varphi(a, y))^{\boldsymbol{M}} = \inf_{b \in A} \varphi^{\boldsymbol{M}}(a, b);$$

(iii) for each multivariable x and quantifiable variable y, there exists a set $\mathcal{S}(x, y)$ of $\mathcal{L}_{\mathcal{C}}(x, y)$-formulas dense in the pseudo-metric space of $\mathcal{L}_{\mathcal{C}}(x, y)$-formulas (with respect to the logical distance), such that for all $\varphi(x, y) \in \mathcal{S}$ and all $a \in A_x$:

$$(\inf_y \varphi(a, y))^{\boldsymbol{M}} = \inf_{b \in A} \varphi^{\boldsymbol{M}}(a, b).$$

Proof. The direction $(i) \Rightarrow (iii)$ follows by applying the definitions, the set $\mathcal{S}$ can be chosen to be $\mathcal{L}_{\mathcal{C}}$. Suppose (iii), and let $\varphi(x, y)$ be an $\mathcal{L}$-formula with y a quantifiable variable. For any $\varepsilon > 0$ we can choose a formula $\psi \in \mathcal{S}(x, y)$ which is in logical distance at most ε from φ. Then

$$(\inf_y \varphi(a, y))^{\boldsymbol{M}} \le (\inf_y \psi(a, y))^{\boldsymbol{M}} + \varepsilon \le \inf_{b \in A} \psi^{\boldsymbol{M}}(a, b) + \varepsilon \le \inf_{b \in A} \varphi^{\boldsymbol{M}}(a, b) + 2\varepsilon.$$

Since ε is arbitrary, this yields $(\inf_y \varphi(a, y))^{\boldsymbol{M}} = \inf_{b \in A} \varphi^{\boldsymbol{M}}(a, b)$, thereby proving (ii).

For the implication $(ii) \Rightarrow (i)$, by letting $\varphi = d(f(x), y)$ and since A is closed, we obtain that A is closed under functions from $\boldsymbol{M}$. Therefore A is the domain of a substructure $\boldsymbol{A}$ of $\boldsymbol{M}$. Continuing we proceed to prove $\boldsymbol{A}$ is elementary in $\boldsymbol{M}$ by induction on the length of formulas. The atomic formulas step, and the non-quantifier connective step are obvious to prove, all that remains is to do the induction step for $\inf_y \varphi(x, y)$ (the sup step is proven by observing $\sup = \neg \inf \neg$). Hence for $a \in A_x$,

$$(\inf_y \varphi(a, y))^{\boldsymbol{M}} = \inf_{b \in A} \varphi^{\boldsymbol{M}}(a, b) = \inf_{b \in A} \varphi^{\boldsymbol{A}}(a, b) = (\inf_y \varphi(a, y))^{\boldsymbol{A}},$$

the second equality is by induction. Therefore $\boldsymbol{A}$ is an elementary substructure of $\boldsymbol{M}$, proving (i). □

Remark 1.30. *The Tarski-Vaught criterion remains valid, with the same proof, even if we replace every occurrence of "structure" with "pre-structure".*

Remark 1.31. *Given that the interpretation of formulas are uniformly continuous, we observe that if a subset $A \subseteq M$ satisfies condition (ii) of the Tarski-Vaught Criterion, then the closure of A in M also satisfies condition (ii). Consequently, the closure of A is the domain of an elementary substructure of $\boldsymbol{M}$.*

Definition 1.32. Let Λ be a non-empty linearly ordered set and $(\boldsymbol{M}_\lambda)_{\lambda\in\Lambda}$ be $\mathcal{L}$-pre-structures such that for any $\lambda \leq \mu$ we have $\boldsymbol{M}_\lambda \preccurlyeq \boldsymbol{M}_\mu$. Then similarly to the construction of the union of an increasing chain of pseudo-metric spaces, in the Preliminaries, we can define the *union of the increasing chain of $\mathcal{L}$-pre-structures* $(\boldsymbol{M}_\lambda)_{\lambda\in\Lambda}$ and denote it by $\bigcup_\lambda \boldsymbol{M}_\lambda$.

Theorem 1.33 (Elementary Chains)**.** *Let Λ be a non-empty linearly ordered set and $(\boldsymbol{M}_\lambda)_{\lambda\in\Lambda}$ be $\mathcal{L}$-structures such that for any $\lambda \leq \mu$ we have $\boldsymbol{M}_\lambda \preccurlyeq \boldsymbol{M}_\mu$. Then $\boldsymbol{M} = \widehat{\bigcup_\lambda \boldsymbol{M}_\lambda}$ is an $\mathcal{L}$-structure such that for any $\lambda \in \Lambda$ we have $\boldsymbol{M}_\lambda \preccurlyeq \boldsymbol{M}$.*

Proof. Obviously $\boldsymbol{N} = \bigcup_\lambda \boldsymbol{M}_\lambda$ is an $\mathcal{L}$-pre-structure. Since $\boldsymbol{N} \preccurlyeq \boldsymbol{M}$, its enough to prove $\boldsymbol{M}_\lambda \preccurlyeq \boldsymbol{N}$ for every λ. This is achieved using the following fact which is proven by induction on the length of $\mathcal{L}$-formulas: for any $\mathcal{L}$-formula $\varphi(x)$, $\lambda \in \Lambda$, and any $a \in (\boldsymbol{M}_\lambda)_x$, $\varphi^{\boldsymbol{N}}(a) = \varphi^{\boldsymbol{M}_\lambda}(a)$. □

Remark 1.34. *Using the notations from Theorem 1.33, if Λ has cofinality $> \aleph_0$ then $\boldsymbol{M} = \bigcup_\lambda \boldsymbol{M}_\lambda$.*

Remark 1.35. *Since our chosen set of connectives $\mathcal{F}$ is countable, the size of the set of $\mathcal{L}$-sentences is equal to $|\mathcal{L}|$. Similarly, the size of all $\mathcal{L}(x)$-formulas where x is a countable multivariable is equal to $|\mathcal{L}|$.*

Definition 1.36. The density character of a metric space X is the smallest cardinality of a dense subset in X, denoted by $\|X\|$. For an $\mathcal{L}$-structure $\boldsymbol{M}$ we will simply use the notation $\|\boldsymbol{M}\|$ for $\|M\|$ and refer to this value as the size of $\boldsymbol{M}$.

The density character of an $\mathcal{L}$-structure has the role in continuous logic that cardinality of a structure has in classical logic.

Theorem 1.37 (Downward Löwenheim-Skolem Theorem)**.** *Let $\boldsymbol{M}$ be an $\mathcal{L}$-structure and a subset $A \subseteq M$. Then $\boldsymbol{M}$ has an elementary substructure $\boldsymbol{N}$ with $A \subseteq N$ and $\|\boldsymbol{N}\| \leq \|A\| + |\mathcal{L}|$.*

Proof. For a nonempty set $B \subseteq M$ we can build a set B' with $|B'| \leq |B| + |\mathcal{L}|$, such that $B \subseteq B'$ and for any $\mathcal{L}$-formula $\varphi(x, y)$ and any $a \in B_x$, the following holds:

$$(\inf_y \varphi(a, y))^{\boldsymbol{M}} = \inf_{b\in B'} \varphi^{\boldsymbol{M}}(a, b).$$

Hence, starting from a dense subset A_0 of A, with cardinality $\|A\|$, we can build a sequence $(A_n)_{n>0}$ such that for any n we have $A_{n+1} = A_n'$. Then $N_0 := \bigcup_n A_n$ has cardinality at most $\|A\| + |\mathcal{L}|$ and satisfies item (ii) from Tarski-Vaught Criteria. Let N be the closure of N_0 in M, and consider any $\mathcal{L}$-formula $\varphi(x, y)$. Then, by Proposition 1.12, for any $a \in N_x$ and for any $\varepsilon > 0$ we can find $\tilde{a} \in (N_0)_x$ such that

$$\begin{aligned}(\inf_y \varphi(a,y))^{\boldsymbol{M}} &\underset{\varepsilon}{\approx} (\inf_y \varphi(\tilde{a},y))^{\boldsymbol{M}} \\ &= \inf_{b\in N_0} \varphi^{\boldsymbol{M}}(\tilde{a},b) \\ &\underset{\varepsilon}{\approx} \inf_{b\in N_0} \varphi^{\boldsymbol{M}}(a,b) \\ &= \inf_{b\in N} \varphi^{\boldsymbol{M}}(a,b),\end{aligned}$$

where $\underset{\varepsilon}{\approx}$ means the numbers are ε close. Sending $\varepsilon \to 0$ we obtain $(\inf_y \varphi(a,y))^{\boldsymbol{M}} = \inf_{b\in N} \varphi^{\boldsymbol{M}}(a,b)$. Applying Tarski-Vaught Criteria yields that N is the domain of an elementary substructure $\boldsymbol{N} \preceq \boldsymbol{M}$ with $\|\boldsymbol{N}\| \leq |N_0| \leq \|A\| + |\mathcal{L}|$. □

Next we present the ultraproduct construction. Let $\mathcal{U}$ be an ultrafilter on a nonempty set I. Let $f : I \to X$, where X is a compact and hausdorff space, be an arbitrary map. Then there is a unique point in X such that the preimages of all its neighborhoods under f are elements of $\mathcal{U}$. This point is denoted by

$$\lim_{\mathcal{U}} f \text{ or } \lim_{\mathcal{U}} f(i) \text{ or } \lim_{I,\mathcal{U}} f.$$

The following observation is very useful: if $U \in \mathcal{U}$, then

$$\lim_{U,\ \mathcal{U}\cap\mathcal{P}(U)} f|_U = \lim_{\mathcal{U}} f,$$

hence restricting the function to an element from the ultrafilter, does not change the limit. This point is denoted by $\lim_U f$ or $\lim_U f(i)$, the ultrafilter being understood from the context. We will make use of the following facts:

(i) if Y is a compact and hausdorff space and $\gamma : X \to Y$ is continuous, then $\lim_{\mathcal{U}} \gamma(f(i)) = \gamma(\lim_{\mathcal{U}} f(i))$;

(ii) if $(S_i)_{i\in I}$ is a family of non-empty subsets of $[0,1]$ then

$$\inf_{a\in\prod S_i} \lim_{\mathcal{U}} a_i = \lim_{\mathcal{U}} \inf S_i.$$

Let $(\boldsymbol{M}_i)_{i\in I}$ be a family of $\mathcal{L}$-pre-structures, and denote $M = \prod M_i$. Then for any two elements $a, b \in M$ we can define $d^{\boldsymbol{M}}(a,b) = \lim_{\mathcal{U}} d^{\boldsymbol{M}_i}(a_i, b_i)$. The map $d^{\boldsymbol{M}} : M \times M \to [0,1]$ is a pseudo-metric.

Proof. It is easy to see that $d^{\boldsymbol{M}}$ is symmetric and $d^{\boldsymbol{M}}(a,a)=0$ for all $a \in M$. It remains to prove the triangle inequality. Let $a,b,c \in M$. For any $\varepsilon > 0$ we can choose an element $U \in \mathcal{U}$ such that the sequences $(d^{\boldsymbol{M}_i}(a_i,b_i))_{i\in U}, (d^{\boldsymbol{M}_i}(b_i,c_i))_{i\in U}$ have the property of being contained in an open interval of diameter $\frac{\varepsilon}{2}$. Therefore we obtain

$$\begin{aligned} d^{\boldsymbol{M}}(a,c) &= \lim_{\mathcal{U}} d^{\boldsymbol{M}_i}(a_i,c_i) = \lim_{U} d^{\boldsymbol{M}_i}(a_i,c_i) \\ &\leq \lim_{U} d^{\boldsymbol{M}_i}(a_i,b_i) + \lim_{U} d^{\boldsymbol{M}_i}(b_i,c_i) + \varepsilon = d^{\boldsymbol{M}}(a,b) + d^{\boldsymbol{M}}(b,c) + \varepsilon. \end{aligned}$$

Now, let $\varepsilon \to 0$. □

The functions on M are defined the usual way, while for relations, e.g. $R \in \mathcal{L}^r$ n-ary, we define $R^{\boldsymbol{M}}(a_1,\ldots,a_n) = \lim_{\mathcal{U}} R^{\boldsymbol{M}_i}((a_1)_i,\ldots,(a_n)_i)$. By using the UC conditions satisfied by each $\boldsymbol{M}_i$, it is easily proven that $\boldsymbol{M} = (M, d^{\boldsymbol{M}}, (R^{\boldsymbol{M}})_{R\in\mathcal{L}^r}, (f^{\boldsymbol{M}})_{f\in\mathcal{L}^f})$ is an $\mathcal{L}$-pre-structure. Therefore we define the *ultraproduct of $(\boldsymbol{M}_i)_{i\in I}$ with respect to $\mathcal{U}$*, also called the *$\mathcal{U}$-ultraproduct of $(\boldsymbol{M}_i)_{i\in I}$*, to be the $\mathcal{L}$-structure $\widehat{\boldsymbol{M}}$, and denote it by $\left(\prod \boldsymbol{M}_i\right)/\mathcal{U}$. We also denote by $\left(\prod_i M_i\right)/\mathcal{U}$ the domain of $\left(\prod_i \boldsymbol{M}_i\right)/\mathcal{U}$. Observe that this construction might collapse more elements in a class than the usual ultraproduct for classical structures would do. If for two elements $a,b \in M$ we have $a_i \neq b_i$ for each $i \in I$ and $d^{\boldsymbol{M}}(a,b)=0$, then a and b would be sent by the natural map $M \to \widehat{M}$ (from Theorem 1.27) to the same element in $\widehat{M}$. The natural morphism $\boldsymbol{M} \to \left(\prod \boldsymbol{M}_i\right)/\mathcal{U}$ is denoted by $a \mapsto [a]$.

Remark 1.38. *In constructing the ultraproduct, if we begin with $\mathcal{L}$-structures rather than $\mathcal{L}$-pre-structures, the process simplifies. Using the same notations, then $(\boldsymbol{M}_0, d^{\boldsymbol{M}_0})$ is already a complete metric space. This allows us to directly define M_0 as the domain of the ultraproduct. Moreover in this case the natural morphism $\boldsymbol{M} \to \left(\prod \boldsymbol{M}_i\right)/\mathcal{U}$ is a projection between the underlying sets.*

Proof. It's enough to prove that the pseudo-metric space $(M, d^{\boldsymbol{M}})$ is complete. For this let (a_n) be a sequence from M, such that $d^{\boldsymbol{M}}(a_n,a_{n+1}) \leq 1/2^n$. Let U_n be a descending chain of sets belonging to $\mathcal{U}$ such that $d^{\boldsymbol{M}_i}((a_k)_i,(a_{k+1})_i) \leq 1/2^k$ for every $i \in U_n$ and $k \leq n$. Then any element $a \in M$ such that a_i is the limit of the cauchy sequence $((a_n)_i)$ if $i \in \bigcap_n U_n$, and $a_i = (a_n)_i$ if $i \in U_n \backslash U_{n+1}$, is a limit of the sequence (a_n). □

Theorem 1.39 (Łoś's Theorem)**.** *Let I be a nonempty set, $(\boldsymbol{M}_i)_{i\in I}$ be a collection of $\mathcal{L}$-pre-structures, $\mathcal{U}$ an ultrafilter on I, and $\boldsymbol{M} = \left(\prod \boldsymbol{M}_i\right)/\mathcal{U}$. Then for every formula $\varphi(x_1,\ldots,x_n)$ and $a^k = (a_i^k)_i \in \prod_{i\in I} M_i$ for $k=1,\ldots,n$, we have*

$$\varphi^{\boldsymbol{M}}([a^1],\ldots,[a^n]) = \lim_{\mathcal{U}} \varphi^{\boldsymbol{M}_i}(a_i^1,\ldots,a_i^n). \tag{1.1}$$

Proof. Let $\varphi = dt_1t_2$ where t_1,t_2 are $\mathcal{L}(x)$-terms. Then by the definition of the ultraproduct we have

$$(dt_1t_2)^{\boldsymbol{M}}([a^1],\ldots,[a^n]) = \lim_{\mathcal{U}} d^{\boldsymbol{M}_i}(t_1^{\boldsymbol{M}_i}(a_i^1,\ldots,a_i^n), t_2^{\boldsymbol{M}_i}(a_i^1,\ldots,a_i^n)).$$

A similar expression also exists for formulas of the type $Rt_1, \ldots, t_m$ where t_i are terms. This proves identity (1.1) for atomic formulas, and now we continue by induction. Let γ any m-ary connective and suppose for $\mathcal{L}$-formulas $\varphi_1, \ldots, \varphi_m$ identity (1.1) holds. Then

$$\begin{aligned}(\gamma\varphi_1, \ldots, \varphi_m)^{\boldsymbol{M}}([a^1], \ldots, [a^n]) &= \gamma(\lim_{\mathcal{U}} \varphi_1^{\boldsymbol{M}_i}(a_i^1, \ldots, a_i^n), \ldots, \lim_{\mathcal{U}} \varphi_m^{\boldsymbol{M}_i}(a_i^1, \ldots, a_i^n)) \\ &= \lim_{\mathcal{U}} \gamma(\varphi_1^{\boldsymbol{M}_i}(a_i^1, \ldots, a_i^n), \ldots, \varphi_m^{\boldsymbol{M}_i}(a_i^1, \ldots, a_i^n)).\end{aligned}$$

While for inf (and similarly for sup), suppose identity (1.1) holds for φ then

$$\begin{aligned}(\inf_x \varphi)^{\boldsymbol{M}}([a^1], \ldots, [a^{n-1}], x) &= \inf_{b \in \prod M_i} \varphi^{\boldsymbol{M}}([a^1], \ldots, [a^{n-1}], [b]) \\ &= \inf_{b \in \prod M_i} \lim_{\mathcal{U}} \varphi^{\boldsymbol{M}_i}(a_i^1, \ldots, a_i^{n-1}, b_i) \\ &= \lim_{\mathcal{U}} (\inf_x \varphi(a_i^1, \ldots, a_i^{n-1}, x))^{\boldsymbol{M}_i},\end{aligned}$$

where the first equality uses the fact that the natural map of $\prod_i M_i \to (\prod_i M_i)/\mathcal{U}$ has dense image, with respect to the metric on the ultrafilter. This completes the induction step, which in turn proves identity (1.1) for all formulas. $\square$

Given an $\mathcal{L}$-structure $\boldsymbol{M}$ and an ultrafilter $\mathcal{U}$ on a nonempty set I we define the $\mathcal{U}$-*ultrapower* by $\boldsymbol{M}^I/\mathcal{U} := (\prod_{i \in I} \boldsymbol{M})/\mathcal{U}$.

Corollary 1.40. *For each $\mathcal{L}$-structure $\boldsymbol{M}$ the diagonal embedding of $\boldsymbol{M}$ into $\boldsymbol{M}^I/\mathcal{U}$ is elementary.*

As a consequence if $\boldsymbol{M}^I/\mathcal{U}$ and $\boldsymbol{N}^I/\mathcal{U}$ are isomorphic then $\boldsymbol{M} \equiv \boldsymbol{N}$. Remarkably the converse holds as well, therefore obtaining an extension of Keisler-Shelah Theorem.

Theorem 1.41. *[GK20, Theorem C.1] For $\boldsymbol{M}$ and $\boldsymbol{N}$ two $\mathcal{L}$-structures, $\boldsymbol{M} \equiv \boldsymbol{N}$ if and only if exists an ultrafilter $\mathcal{U}$ on a nonempty set I such that $\boldsymbol{M}^I/\mathcal{U} \cong \boldsymbol{N}^I/\mathcal{U}$.*

Definition 1.42. Let Σ be a set of closed conditions. We say Σ is *finitely satisfiable* if any finite subset of Σ is satisfiable. We also say Σ is *approximately finitely satisfiable* if for any finite subset $\{\sigma_1 = 0, \ldots, \sigma_n = 0\} \subseteq \Sigma$ and any $\varepsilon > 0$ the set $\{\sigma_1 \leq \varepsilon, \ldots, \sigma_n \leq \varepsilon\}$ is satisfiable.

Theorem 1.43 (Compactness Theorem for Continuous Logic)**.** *Let Σ be a family of closed $\mathcal{L}$-conditions. Then the following are equivalent:*

(i) Σ *is satisfiable;*

(ii) Σ *is finitely satisfiable;*

(iii) Σ *is approximately finitely satisfiable.*

Proof. The directions $(i) \Rightarrow (ii)$ and $(ii) \Rightarrow (iii)$ are trivial. Hence, suppose Σ is approximately finitely satisfiable. Let

$$\Lambda = \{(\lambda, \varepsilon) : \varepsilon > 0 \text{ and } \lambda \text{ a finite subset of } \Sigma\}.$$

Then for each $(\lambda, \varepsilon) \in \Lambda$ we can pick an $\mathcal{L}$-structure $\boldsymbol{M}_{\lambda,\varepsilon}$ such that the corresponding truth value of all sentences from the conditions of λ are smaller than ε. For each condition $\sigma = 0 \in \Sigma$ and each $\varepsilon > 0$ consider the set

$$T_{\sigma,\varepsilon} = \{(\lambda, \delta) \in \Lambda : \sigma = 0 \in \lambda \text{ and } \delta \leq \varepsilon\}.$$

Then $\{T_{\sigma,\varepsilon}\}_{\sigma=0\in\Sigma,\varepsilon>0}$ is a family on $\mathcal{P}(\Lambda)$ with the finite intersection property, hence it can be extended to an ultrafilter $\mathcal{U}$. All that remains is to apply Łoś's Theorem to $(\prod_{(\lambda,\varepsilon)\in\Lambda} \boldsymbol{M}_{\lambda,\varepsilon})/\mathcal{U}$, and to observe that this ultraproduct satisfies Σ. □

As a consequence of condition (iii) above, we obtain:

Corollary 1.44. *Let Σ be a family of closed $\mathcal{L}$-conditions and φ be an $\mathcal{L}$-sentence. Then the following are equivalent:*

(i) $\Sigma \models \varphi = 0$*;*

(ii) for every $\varepsilon > 0$ exists $\delta > 0$ and a finite subset of conditions $\{\varphi_1 = 0, \dots, \varphi_n = 0\}$ of Σ such that
$$\{\varphi_1 \leq \delta, \dots, \varphi_n \leq \delta\} \models \varphi \leq \varepsilon.$$

Proof. The direction $(ii) \to (i)$ is trivial. Suppose (i) and let $\varepsilon > 0$. Then the family of conditions $\Sigma \cup \{\varepsilon \dot{-} \varphi = 0\}$ is not satisfiable, and by compactness is not approximately finitely satisfiable. Hence there exists $\delta > 0$ and $\{\varphi_1 = 0, \dots, \varphi_n = 0\} \subseteq \Sigma$ such that

$$\{\varphi_1 \leq \delta, \dots, \varphi_n \leq \delta, \varepsilon \dot{-} \varphi \leq \delta\}$$

is not satisfiable. Therefore, for an $\mathcal{L}$-structure $\boldsymbol{M}$, if $\boldsymbol{M} \models \varphi_1 \leq \delta, \dots, \varphi_n \leq \delta$ then $(\varepsilon \dot{-} \varphi)^{\boldsymbol{M}} > \delta$, in particular $\boldsymbol{M} \models \varphi \leq \varepsilon$. □

Remark 1.45. *Unfortunately, not every basic result in first-order classical logic translates immediately to continuous logic. For instance, if (T_n) is an increasing chain of theories then it is not necessarily true that $T = \bigcup T_n$ is a theory. As an example, consider $T_n = \mathrm{Th}\left(\{\sigma \leq \frac{1}{m+1} : m \leq n\}\right)$ for some sentence σ.*

Theorem 1.46 (Upward Löwenheim-Skolem Theorem)**.** *Let $\boldsymbol{M}$ be a non-compact $\mathcal{L}$-structure. Then, for every cardinal κ, $\boldsymbol{M}$ has an elementary extension of size $\geq \kappa$ (as in Definition 1.36).*

Proof. A metric space is compact if and only if it is totally bounded and complete [Roy10, Theorem 16, p. 199]. Therefore exists an $\varepsilon > 0$ such that for every n exists a sequence

$a_1, \dots, a_n \in M$ with $d^{\boldsymbol{M}}(a_i, a_j) \geq \varepsilon$ for every $1 \leq i, j \leq n$ with $i \neq j$. Let $\{c_i : i < \kappa\}$ be constant symbols. Then

$$\mathrm{Th}(\boldsymbol{M}_M) \cup \{d^{\boldsymbol{M}}(c_i, c_j) \geq \varepsilon : i, j < \kappa, i \neq j\}$$

is finitely satisfiable, and by compactness has a model $\boldsymbol{N}$, which is an elementary extension of $\boldsymbol{M}$ with $\|\boldsymbol{N}\| \geq \kappa$. □

Unnested Formulas

Throughout this subsection x is a finite multivariable.

Definition 1.47. Let φ be an $\mathcal{L}(x)$-formula. We say φ is *unnested* if all its atomic subformulas are of the form:

(i) $Rx_1 \dots x_n$ where $R \in \mathcal{L}^r$ and $x_1, \dots, x_m$ are distinct variables;

(ii) $d(f(x_1, \dots, x_m), y)$ where $f \in \mathcal{L}^f$ and $x_1, \dots, x_m, y$ are distinct variables.

(iii) $d(x, y)$ where x, y are distinct variables.

Moreover, we say φ is *approximable by unnested formulas* if for every $\varepsilon > 0$ there exists an unnested formula φ_ε such that $\models \sup_x |\varphi - \varphi_\varepsilon| \leq \varepsilon$.

Proposition 1.48. *Every $\mathcal{L}$-formula $\varphi(x)$ is approximable by unnested formulas.*

Proof. Since formulas are uniformly continuous it's enough to prove that every atomic formula is approximable by unnested formulas. This is done by an induction on the complexity of the terms appearing in the atomic formulas.

Fix $\varepsilon > 0$ and let $y, y_1, \dots, y_n$ be distinct single variables not appearing in x. Let α be an atomic formula. Suppose $\alpha(x, y) = d(t, y)$ with $t = ft_1 \dots t_n$, where $t, t_1, \dots, t_n$ are $\mathcal{L}(x)$-terms. Since f has a modulus of continuity results, that there exists $\delta > 0$ such that for any model $\boldsymbol{M}$, $a, b \in M^n$, we have $d(a, b) \leq \delta \implies |d(f(a), f(b))| \leq \varepsilon$. Then

$$\models \sup_{xy} \left| \sup_{y_1,\dots,y_n} \left(\frac{\delta - (d(t_1, y_1) + \dots + d(t_n, y_n))}{\delta} d(f(y_1 \dots y_n), y) \right) - d(t, y) \right| \leq \varepsilon$$

Since $d(t_m, y_m)$ have smaller term complexity, by the induction hypothesis and the fact that formulas in continuous logic are uniformly continuous, results that there exists $\delta' > 0$ and unnested formulas $\varphi_1, \dots, \varphi_n$ such that

$$\models \sup_{xy_1} |d(t_1, y_1) - \varphi_1| \leq \delta', \dots, \sup_{xy_n} |d(t_n, y_n) - \varphi_n| \leq \delta' \text{ and}$$

$$\models \sup_{xy} \left| \sup_{y_1,\dots,y_n} \left(\frac{\delta - (\varphi_1 + \dots + \varphi_n)}{\delta} d(f(y_1 \dots y_n), y) \right) - d(t, y) \right| \leq 2\varepsilon$$

we observe that $\sup_{y_1,\dots,y_n} \frac{\delta - (\varphi_1 + \dots + \varphi_n)}{\delta} d(f(y_1 \dots y_n), y)$ is an unnested formula which represents a 2ε approximation to α. Therefore α is approximable by unnested formulas.

Now, suppose $\alpha(x) = Rt_1, \dots, t_n$ (similarly for dt_1t_2) where $t_1, \dots, t_n$ are $\mathcal{L}(x)$-terms. Fix $\varepsilon > 0$. Since R has a modulus of continuity, there exists $\delta > 0$ such that for any model $\boldsymbol{M}$, $a, b \in M^n$, we have $d(a,b) \leq \delta \implies |R(a) - R(b)| \leq \varepsilon$. Then

$$\models \sup_x \left| \sup_{y_1,\dots,y_n} \left(\frac{\delta - (d(t_1,y_1) + \cdots + d(t_n,y_n))}{\delta} R(y_1,\dots,y_n) \right) - R(t_1,\dots,t_n) \right| \leq \varepsilon.$$

By the previous paragraph, for each m, the term $d(t_m, y_m)$ is approximable by unnested formulas. Therefore, by applying the same reasoning, we conclude that α is uniformly approximable by unnested formulas. □

1.4 Type Spaces

Let $\mathcal{L}$ be a language, $x = (x_i)_{i \in I}$ be a multivariable, $\boldsymbol{M}$ be an $\mathcal{L}$-structure, and p be a set of $\mathcal{L}(x)$-conditions. We call p an *x-type* if there is an $\mathcal{L}$-structure $\boldsymbol{N}$ and $a \in N_x$ such that $\boldsymbol{N} \models E(a)$ for every $E(x) \in p$. Without referring to the structure we call a a *realization* of p, and denote this by $a \models p$. By compactness this is also equivalent to p being finitely satisfiable or p being approximately finitely satisfiable. A maximal satisfiable set of $\mathcal{L}(x)$-conditions is called a *complete* type.

Let Σ be a set of closed conditions. We say that p is *Σ-realizable* if $p \cup \Sigma$ is a type.

Definition 1.49. The set of all Σ-realizable complete x-types is denoted by $\mathrm{S}_x(\Sigma)$. Let A be a subset of M, and q be a set of $\mathcal{L}_A(x)$-conditions. We call q an *x-type over A in $\boldsymbol{M}$* if it is $\mathrm{Th}(\boldsymbol{M}_A)$-realizable. The set of all $\mathrm{Th}(\boldsymbol{M}_A)$-realizable complete x-types is denoted by $\mathrm{S}_x(A)$.

Remark 1.50. *Even though these definitions are very similar to classical logic, there is one striking difference: for $E(x)$ an $\mathcal{L}(x)$-condition,*

$$\boldsymbol{M} \preccurlyeq \boldsymbol{N} \not\Longrightarrow (\boldsymbol{N} \text{ realizes } E(x) \text{ implies } \boldsymbol{M} \text{ realizes } E(x)),$$

while is true that,

$$\boldsymbol{M} \preccurlyeq \boldsymbol{N} \implies (\boldsymbol{N} \text{ realizes } E(x) \text{ implies } \boldsymbol{M} \models \inf_x E(x) = 0).$$

Therefore we arrive at the following:

Definition 1.51. We say that p is *approximately finitely satisfiable in $\boldsymbol{M}$* if for every finite subset $\{\varphi_1(x) = 0, \dots, \varphi_n(x) = 0\} \subseteq p$ and every $\varepsilon > 0$ the set $\{\varphi_1(x) \leq \varepsilon, \dots, \varphi_n(x) \leq \varepsilon\}$ is satisfiable in $\boldsymbol{M}$.

Proposition 1.52. *Suppose p is a set of $\mathcal{L}_A(x)$-conditions. The following are equivalent:*

(i) p is an x-type over A in $\boldsymbol{M}$,

(ii) p is approximately finitely satisfiable in $\boldsymbol{M}$,

(iii) for all conditions $\varphi_1(x) = 0, \ldots, \varphi_n(x) = 0$ in p, we have

$$\inf_x \max(\varphi_1, \ldots, \varphi_n) = 0$$

belongs to $\mathrm{Th}(\boldsymbol{M}_A)$.

Remark 1.53. *Let p be a complete x-type and $a \in M_x$ a realization of p. Then*

$$q = \{E(x) : E(x) \text{ an } \mathcal{L}_{\mathcal{C}}(x)\text{-condition and } \boldsymbol{M} \models E(a)\}$$

is a complete x-type with conditions from $\mathcal{L}_{\mathcal{C}}(x)$, extending p. Applying Theorem 1.15, we conclude that q is the unique complete x-type over $\mathcal{L}_{\mathcal{C}}(x)$-conditions extending p. Thus, restricting $\mathcal{C}$ to any full system of connectives does not change the set of complete types.

Definition 1.54. For $a \in M_x$ we define the *type of a in $\boldsymbol{M}$* by

$$\mathrm{tp}^{\boldsymbol{M}}(a) = \{E(x) : E(x) \text{ an } \mathcal{L}(x)\text{-condition and } \boldsymbol{M} \models E(a)\},$$

and the *type of a over a subset $A \subseteq M$* by $\mathrm{tp}^{\boldsymbol{M}}(a/A) = \mathrm{tp}^{\boldsymbol{M}_A}(a)$.

Definition 1.55. We say that the $\mathcal{L}(x)$-formulas φ and ψ are Σ-*equivalent*, denoted by $\varphi \equiv_\Sigma \psi$, if $\Sigma \models \sup_x |\varphi(x) - \psi(x)| = 0$. This is an equivalence relation on the set of $\mathcal{L}(x)$-formulas and we define $\mathrm{B}_x(\Sigma) = \mathcal{L}(x)/\equiv_\Sigma$.

Remark 1.56. *For an $\mathcal{L}$-formula $\varphi(x)$, let $[\varphi]$ denote its equivalence class in $\mathrm{B}_x(\Sigma)$. Then for any type $p \in \mathrm{S}_x(\Sigma)$ if $\varphi \in p$ then $[\varphi] \subseteq p$. Therefore p is closed under the relation $\equiv_\Sigma$, and if we denote by $[p]$ the classes of all the formulas in p then $[p] \in \mathcal{P}(\mathrm{B}_x(\Sigma))$. Moving forward we will have this perspective in mind, and the aim of checking, in regards to types, that everything is well-defined up to Σ-equivalence.*

For p in $\mathrm{S}_x(\Sigma)$ and φ an $\mathcal{L}(x)$-formula we observe that there exists an element $t \in [0,1]$ such that $\{\varphi \leq r : t < r\} \cup \{\varphi \geq r : t > r\} \subseteq p$, then for every realization $a \models p$ with $a \in M_x$, for some structure $\boldsymbol{M}$, we have $t = \varphi^{\boldsymbol{M}}(a)$. Since this value is independent of the realization of the type we denote $\varphi^p := t$. Moreover this is well-defined up to Σ-equivalence. Therefore:

Definition 1.57. Given an $\mathcal{L}(x)$-formula φ we define $\dot{\varphi} : \mathrm{S}_x(\Sigma) \to [0,1]$ by $\dot{\varphi}(p) = \varphi^p$.

Remark 1.58. *For φ, ψ two $\mathcal{L}(x)$-formulas, $\dot{\varphi} = \dot{\psi}$ if and only if $\varphi \equiv_\Sigma \psi$.*

We refer to the minimal topology on $\mathrm{S}_x(\Sigma)$ that makes the functions $(\dot{\varphi})_{\varphi \in \mathcal{L}(x)}$ continuous as the *logic topology*.

Theorem 1.59. *The space $\mathrm{S}_x(\Sigma)$, endowed with the logic topology, is compact and Hausdorff.*

Proof. By the properties of the product topology, the function

$$\zeta : \mathrm{S}_x(\Sigma) \to [0,1]^{\mathcal{L}(x)},\ p \to (\dot{\varphi}(p))_{\varphi\in\mathcal{L}(x)},$$

is a homeomorphism onto its image $\zeta(\mathrm{S}_x(\Sigma))$. By Tychonoff's theorem $[0,1]^{\mathcal{L}(x)}$ is compact, hence all that remains is to prove that $\zeta(\mathrm{S}_x(\Sigma))$ is closed. Let f be in the closure of $\zeta(\mathrm{S}_x(\Sigma))$. For all $\mathcal{L}(x)$-formulas $\varphi_1, \dots, \varphi_n$ and any $\varepsilon > 0$ there exists a type $p \in \mathrm{S}_x(\Sigma)$ such that $|\varphi_i^p - f(\varphi_i)| \leq \varepsilon$ for all $i \leq n$. But this is equivalent to the set of conditions $Q = \Sigma \cup \{\varphi(x) = f(\varphi) : \varphi \in \mathcal{L}(x)\}$ being approximately finitely satisfiable. By compactness Q is satisfiable and there is a unique $q \in \mathrm{S}_x(\Sigma)$ such that $Q \subseteq q$. Then $\zeta(q) = f$, implying $f \in \zeta(\mathrm{S}_x(\Sigma))$. Hence $\zeta(\mathrm{S}_x(\Sigma))$ is closed. □

Remark 1.60. *Observe that if $\varphi(x)$ is an $\mathcal{L}_{\mathcal{C}}(x)$-formula, using the Remark 1.53, it is straightforward to define $\dot{\varphi}$, as in Definition 1.57. Then, by Proposition 1.15, there exists a sequence (φ_n) of $\mathcal{L}(x)$-formulas such that $(\dot{\varphi}_n)$ converges uniformly to $\dot{\varphi}$. Since $(\dot{\varphi}_n)$ is a sequence of continuous functions, $\dot{\varphi}$ is also a continuous function. Therefore the logic topology is the same if we define it with any other full system of connectives, since they all are equal to the logic topology when we allow all possible connectives $\mathcal{C}$.*

Proposition 3.4 from [BYU10] provides an important result:

Proposition 1.61. *The set $\{\dot{\varphi} : \varphi \in \mathcal{L}(x)\}$ is dense in the set of continous functions* $\mathrm{C}(\mathrm{S}_x(\Sigma), [0,1])$ *with respect to the uniform metric.*

Proof. This is a consequence of Proposition 1.4. □

Remark 1.62. *Using Proposition* 1.15*, for any full system of connectives $\mathcal{G}$, the set $\{\dot{\varphi} : \varphi \in \mathcal{L}_{\mathcal{G}}(x)\}$ is dense in* $\mathrm{C}(\mathrm{S}_x(\Sigma), [0,1])$ *with respect to the uniform metric.*

For any $\varphi(x)$ and $\varepsilon \in (0,1]$ we define $[\varphi < \varepsilon] = \{p \in \mathrm{S}_x(\Sigma) : \varphi^p < \varepsilon\}$ and $[\varphi > \varepsilon] = \{p \in \mathrm{S}_x(\Sigma) : \varphi^p > \varepsilon\}$. In a similar way we define $[\varphi \leq \varepsilon]$ and $[\varphi \geq \varepsilon]$. It is immediate that $[\varphi > \varepsilon] = [\neg\varphi < 1 - \varepsilon]$ and that the sets $[\varphi < \varepsilon]$ are open in $\mathrm{S}_x(\Sigma)$. Moreover, given $\delta \in (0,1]$ and $p \in [\varphi < \varepsilon] \cap [\psi < \delta]$, there exist dyadic reals $e, d \in [0,1]$, and $r \in (0,1]$, with $\varphi^p < e < \varepsilon, \psi^p < d < \delta$ such that

$$p \in [\max(|\varphi - e|, |\psi - d|) < r] \subseteq [\varphi < \varepsilon] \cap [\psi < \delta].$$

Finally, as a consequence of these observations we obtain:

Proposition 1.63. *The collection $\{[\varphi < \varepsilon] : \varepsilon \in (0,1], \varphi \in \mathcal{L}(x)\}$ forms a basis of the logic topology on* $\mathrm{S}_x(\Sigma)$.

Corollary 1.64. *For a multivariable $y \subseteq x$, the map $\eta_{x,y} : \mathrm{S}_x(\Sigma) \to \mathrm{S}_y(\Sigma)$ defined by*

$$p \mapsto p|_y := \{E(y) : E \textit{ an } \mathcal{L}(y)\textit{-condition and } E(y) \in p\},$$

is a continuous surjection.

The following proposition establishes the relationship between the functions $\dot{\varphi}$ and the logical connectives.

Proposition 1.65. *Let γ be an n-ary connective, $\varphi_1, \ldots, \varphi_n$ be $\mathcal{L}(x)$-formulas, and $\varphi(x, y)$ be an $\mathcal{L}(xy)$-formula with y a quantifiable variable. Then the following hold:*

(i) $(\gamma\varphi_1 \dot{\ldots} \varphi_n) = \gamma(\dot{\varphi}_1, \ldots, \dot{\varphi}_n)$;

(ii) for every $q \in \mathrm{S}_x(\Sigma)$ we have $(\dot{\inf_y} \varphi)(q) = \inf\{\dot{\varphi}(p) : q \subseteq p, p \in \mathrm{S}_{xy}(\Sigma)\}$ or more succintly

$$(\dot{\inf_y} \varphi)(q) = \inf_{q \subseteq p} \dot{\varphi}(p).$$

Moreover exists $p \in \mathrm{S}_{xy}(\Sigma)$, with $q \subseteq p$, such that $(\dot{\inf_y} \varphi)(q) = \dot{\varphi}(p)$

Proof. Item (i) is straightforward. For item (ii), let $q \in \mathrm{S}_x(\Sigma)$. For every $p \in \mathrm{S}_{xy}(\Sigma)$ with $q \subseteq p$, by taking a realization of p, we observe that $\dot{\varphi}(p) \geq (\dot{\inf_y} \varphi)(q)$. Consider the set of conditions $q \cup \{\varphi(xy) = (\dot{\inf_y} \varphi)(q)\}$, which is approximately finitely satisfiable, hence it is xy-type. Then there exists a complete xy-type p, containing q, such that $\dot{\varphi}(p) = (\dot{\inf_y} \varphi)(q)$, from which immediately results $(\dot{\inf_y} \varphi)(q) = \inf_{q \subseteq p} \dot{\varphi}(p)$. □

Remark 1.66. *Item (ii) of Proposition 1.65, in a straightforward way generalizes to a similar statement with the change that y is a finite multivariable of quantifiable variables. Moreover, it can also be generalized to a situation, using Remark 1.9, when y is any quantifiable multivariable.*

For the remainder of this section, let κ denote a cardinal, and v a single variable of $\mathcal{L}$.

Definition 1.67. An $\mathcal{L}$-structure $\boldsymbol{M}$ is said to be *κ-saturated* if for all $A \subseteq M$ of size $< \kappa$, that is $|A| < \kappa$, each v-type over A in $\boldsymbol{M}$ is realized in $\boldsymbol{M}$.

Remark 1.68. *Suppose $\boldsymbol{M}$ is 1-saturated, and $\varphi \in \mathcal{L}(v)$. Then, $\boldsymbol{M} \models \inf_v \varphi(v) = 0$ is equivalent to $\{\varphi(v) \leq 1/n : n \geq 1\}$ being finitely satisfiable in $\boldsymbol{M}$, which in turn by saturation is equivalent to the existence of some element $a \in M$ such that $\boldsymbol{M} \models \varphi(a)$. This interesting phenomenon is captured by Proposition 1.71.*

Remark 1.69. *In Definition 1.67, we use cardinality as the notion of size for A. On the other hand according to Definition 1.36, in continuous logic the appropriate notion of size is the cardinality of a minimal dense set. In this context, using the uniform continuity of formulas it is easy to see that if a model is κ-saturated then it also realizes every type over A with $\|A\| < \kappa$. Therefore, Definition 1.67 remains equivalent regardless of whether size is interpreted as $|A|$ or $\|A\|$.*

Definition 1.70. For $Q_i \in \{\exists, \forall\}$ and $\varphi \in \mathcal{L}(x_1, \ldots, x_n)$ we denote by

$$\boldsymbol{M} \models Q_1 x_1 \ldots Q_n x_n \varphi(x_1, \ldots, x_n) = 0$$

if

$$(Q_1 a_1 \in M) \ldots (Q_n a_n \in M)(\boldsymbol{M} \models \varphi(a_1, \ldots, a_n) = 0).$$

We write $\dot{\exists} := \inf$ and $\dot{\forall} = \sup$.

Proposition 1.71. *Suppose $\boldsymbol{M}$ is an $\aleph_0$-saturated $\mathcal{L}$-structure and $\varphi \in \mathcal{L}_M(x)$ with $x = (x_1, \ldots, x_n)$. Then*

$$\boldsymbol{M} \models \dot{Q}_1 x_1 \ldots \dot{Q}_n x_n \varphi(x) = 0 \Longleftrightarrow \boldsymbol{M} \models Q_1 x_1 \ldots Q_n x_n \varphi(x) = 0$$

This is proved by induction on n. Therefore, for a κ-saturated model, in certain situations we can safely exchange the continuous quantifiers with the classical ones. It is evident that $\boldsymbol{M}$ is κ-saturated if and only if for all $A \subseteq M$ of size $< \kappa$, each $p \in \mathrm{S}_v(A)$ is realized in $\boldsymbol{M}$. In fact an even stronger result holds:

Proposition 1.72. *Suppose κ is an infinite cardinal, $\boldsymbol{M}$ is κ-saturated, $A \subseteq M$ of size $< \kappa$, and $|x| \leq \kappa$. Then every $p \in \mathrm{S}_x(A)$ is realized in $\boldsymbol{M}$.*

Proof. We can assume $|x| = \kappa$ and index $x = (x_\delta)_{\delta<\kappa}$. For $\delta < \kappa$ we denote

$$p_\delta = \{\varphi = 0 \in p : \varphi(x) \in \mathcal{L}((x_i)_{i \leq \delta})\}.$$

We will construct recursively a sequence $(a_\delta)_{\delta<\kappa} \in M_x$ such that $(a_i)_{i\leq\delta} \models p_\delta$ for every $\delta < \kappa$. By κ-saturation we can pick an $a_0 \in M$ such that $a_0 \models p_0$. Let $0 < \delta < \kappa$ and suppose we constructed the sequence $(a_i)_{i<\delta}$. Let $q = \{\varphi((a_i)_{i<\delta}, x_\delta) = 0 : \varphi = 0 \in p_\delta\}$ and consider a finite set $\{\varphi_j((a_i)_{i<\delta}, x_\delta) = 0 : j = 1, \ldots, n\} \subseteq q$. Then the $\mathcal{L}((x_i)_{i<\delta})$-condition $\inf_{x_\delta} \max(\varphi_1 \ldots \varphi_n) = 0$ belongs to p, and by recursion

$$\boldsymbol{M} \models (\inf_{x_\delta} \max(\varphi_1 \ldots \varphi_n) = 0)(a_i)_{i<\delta}.$$

Applying Proposition 1.71, yields an a in M such that $\boldsymbol{M} \models \varphi_j((a_i)_{i<\delta}, a) = 0$ for $j = 1, \ldots, n$. Therefore q is an x_δ-type over $A \cup \{a_i : i < \delta\}$. By κ-saturation we can choose $a_\delta \in M$ such that $a_\delta \models q$, and then obviously $(a_i)_{i\leq\delta} \models p_\delta$. Finally, we obtain that $(a_i)_{i<\kappa} \models p$. □

Remark 1.73. *Observe that in the proof of Theorem 1.72 it was not necessary to appeal to Proposition 1.71 to prove q is an x_δ-type over $A \cup \{a_i : i < \delta\}$. It is enough to observe that it is approximately finitely satisfiable and then by Proposition 1.52 it's a type.*

Theorem 1.74. *Every $\mathcal{L}$-structure $\boldsymbol{M}$ has a κ-saturated elementary extension.*

Proof. First we will prove the claim: every $\mathcal{L}$-structure $\boldsymbol{M}$ has an elementary extension $\boldsymbol{N}$ such that all $p \in \mathrm{S}_v(M)$ are realized in $\boldsymbol{N}$.

Proof of claim. For every $p \in \mathrm{S}_v(M)$ we assign a single variable v_p, such that $v_p \neq v_q$ for all $p \neq q$. Let $\dot{p}$ be the set of conditions obtained by the substitution of v with v_p in the conditions of p. Clearly $\dot{p}$ belongs to $\mathrm{S}_{v_p}(M)$. Since the set $\bigcup_{p\in\mathrm{S}_v(M)} \dot{p}$ is finitely satisfiable, applying compactness, we obtain the required object $\boldsymbol{N}$. □

Back to the theorem, we can suppose κ is regular. Then starting from $\boldsymbol{M}_0 = \boldsymbol{M}$ we can build a elementary chain $(\boldsymbol{M}_\delta)_{\delta<\kappa}$ such that for all $\delta < \kappa, \boldsymbol{M}_{\delta+1}$ realizes every type in $\mathrm{S}_v(\boldsymbol{M}_\delta)$. Let $\boldsymbol{P} = \bigcup_{\delta<\kappa} \boldsymbol{M}_\delta$. Then, by Proposition 1.33, $\boldsymbol{M} \preccurlyeq \boldsymbol{P}$. Let $A \subseteq P$, with $|A| < \kappa$, and $p \in \mathrm{S}_v(A)$. Then exists $\delta < \kappa$ such that $A \subseteq M_\delta$. Therefore, by construction, $p \in \mathrm{S}_v(M_\delta)$, which is realized in $\boldsymbol{M}_{\delta+1}$, hence realized in $\boldsymbol{P}$. □

In order to define strongly κ-homogeneous structures we need to generalize the notion of an elementary map.

Definition 1.75. Let $\boldsymbol{M}, \boldsymbol{N}$ be $\mathcal{L}$-structures with parameter sets $A \subseteq M, B \subseteq N$. We call $f : A \to B$ an *elementary map* if for all $\mathcal{L}$-formulas $\varphi(x)$ we have $\varphi^{\boldsymbol{M}}(a) = \varphi^{\boldsymbol{N}}(f(a))$ for all $a \in A_x$. In order to emphasize the structures, we say f is an elementary map with respect to $\boldsymbol{M}, \boldsymbol{N}$, and if $\boldsymbol{M} = \boldsymbol{N}$ we simply say f is elementary with respect to $\boldsymbol{M}$.

If f is elementary, then f is injective, $f^{-1} : f(A) \to M$ is elementary, and $\boldsymbol{M}_A \equiv \boldsymbol{N}_{f(A)}$, where $\boldsymbol{N}_{f(A)}$ is the expansion of $\boldsymbol{N}$ to an $\mathcal{L}_A$-structure where each name for a in A is interpreted by the element $f(a)$ of N. Moreover, if $f : M \to M$ is elementary and surjective, then it is an automorphism of $\boldsymbol{M}$.

Definition 1.76. We call $\boldsymbol{M}$ *strongly κ-homogeneous* if for any $A \subseteq M$ of size $< \kappa$, any elementary map $f : A \to M$ can be extended to an automorphism of $\boldsymbol{M}$.

Theorem 1.77. *Suppose κ is infinite. Then $\boldsymbol{M}$ has a κ-saturated elementary extension $\boldsymbol{N}$, such that each reduct of $\boldsymbol{N}$ is strongly κ-homogeneous.*

Proof. We can suppose κ is regular. Let $\boldsymbol{M}_0 = \boldsymbol{M}$ and we build an elementary chain $(\boldsymbol{M}_\delta)_{\delta<\kappa}$ with the property that $\boldsymbol{M}_{\delta+1}$ is $|M_\delta|^+$-saturated. Then $\boldsymbol{N} = \bigcup_\delta \boldsymbol{M}_\delta$ is the required object. This is a consequence of the fact that by a recursive construction we can extend an elementary map $f : A \to N$, with $A \subseteq N$ of size $< \kappa$, to an automorphism of $\boldsymbol{N}$. The idea of the recursion is conveyed by the following method of extending the domain of f by an element $a \in N \backslash A$. Let $p = \mathrm{tp}^{\boldsymbol{N}}(a/A)$, a v-type, and $\delta < \kappa$ such that $A \cup f(A) \subseteq M_\delta$. Since the map f is also elementary with respect to $\boldsymbol{M}_\delta$, and p is a type in $\boldsymbol{M}_\delta$, $f(p)$ is a v-type in $\boldsymbol{M}_\delta$. But then, the $|M_\delta|^+$ saturation of $\boldsymbol{M}_{\delta+1}$ yields the existence of $a' \in \boldsymbol{M}_{\delta+1}$ such that $a' \models f(p)$. All that remains is to see that $f \cup \{(a, a')\}$ is an elementary map with respect to the structure $\boldsymbol{M}_{\delta+1}$, and therefore also an elementary map with respect to $\boldsymbol{N}$.

For a reduct $\mathcal{L}' \subseteq \mathcal{L}$, if we consider $\boldsymbol{M}_\delta|_{\mathcal{L}'}$ the $\mathcal{L}'$ reduct of $\boldsymbol{M}_\delta$ and $\boldsymbol{N}|_{\mathcal{L}'}$ the $\mathcal{L}'$ reduct of $\boldsymbol{N}$, then the exact same proof proves $\boldsymbol{N}|_{\mathcal{L}'}$ is strongly κ-homogeneous. $\square$

Definition 1.78. We call a κ-saturated and strongly κ-homogeneous model of a complete theory T a *κ-monster model of T*, or simply *monster model of T* under the assumption that κ is chosen to be sufficiently large for our purposes.

Let T be a complete theory, and let $x = (x_i)_{i\in I}, y = (y_i)_{i\in I}$ be two disjoint multivariables. Using Theorem 1.74 and the completeness of T, we can choose a model $\boldsymbol{M}$ of T that realizes every type in $\mathrm{S}_{xy}(T)$. For any $p, q \in \mathrm{S}_x(T)$ we define *the distance between p, q* to be

$$d(p, q) = \inf\{d^{\boldsymbol{M}}(a, b) : a, b \in M_x, a \models p, b \models q\}.$$

If $\boldsymbol{N}$ is a model of T which realizes every type in $\mathrm{S}_{xy}(T)$, and $a, b \in M_x$ such that $a \models p, b \models q$, then by considering the xy-type $\mathrm{tp}^{\boldsymbol{M}}(ab)$, we obtain $a', b' \in N_x$ such that $a' \models p, b' \models q$ and $d^{\boldsymbol{N}}(a', b') = d^{\boldsymbol{M}}(a, b)$. The last observation implies that the distance between p, q is independent of the chosen model of T, and therefore we obtain a map $d : \mathrm{S}_x(T) \times \mathrm{S}_x(T) \to [0,1]$.

We note that $d(p,p) = 0$ and $d(p,q) = d(q,p)$. Furthermore, by a straightforward argument with $\boldsymbol{M}$ as above, we can find $a, b \in M_x$ such that $d(p,q) = d^{\boldsymbol{M}}(a,b)$. Hence, if $d(p,q) = 0$ then $p = q$. To complete the proof that d is a metric, it remains to prove the triangle inequality:

Proof. We choose $\boldsymbol{M}$ to be an $\max(\aleph_0, |x|^+)$-monster model of T. Let $p, q, r \in \mathrm{S}_x(T)$ and $a \models p, b \models q, b' \models q$ and $c \models r$. Then by the homogeneity of $\boldsymbol{M}$ exists an automorphism h of $\boldsymbol{M}$ such that $h(b) = b'$. Therefore

$$d(p,r) \leq d^{\boldsymbol{M}}(a, h^{-1}(c)) \leq d^{\boldsymbol{M}}(a,b) + d^{\boldsymbol{M}}(b, h^{-1}(c)) = d^{\boldsymbol{M}}(a,b) + d^{\boldsymbol{M}}(b',c).$$

Since a, b, b', c are arbitrary realizations of the respective types, we get $d(p,r) \leq d(p,q) + d(q,r)$. □

Definition 1.79. The topology on $\mathrm{S}_x(T)$ induced by the metric d is referred to as the *metric topology.*

Using Proposition 1.12, it is straightfoward to verify that the metric topology is finer than the logic topology on $\mathrm{S}_x(T)$. Moreover:

Theorem 1.80. *The metric space* $(\mathrm{S}_x(T), d)$ *is complete.*

Proof. Let (p_n) be a cauchy sequence in $\mathrm{S}_x(T)$ and $\boldsymbol{M} \models T$ realize every type in $\mathrm{S}_{xy}(T)$. Let (q_n) be a subsequence of (p_n) such that $d(q_n, q_{n+1}) \leq \frac{1}{2^n}$. Choose a sequence (a_n) in M_x, such that for every n we have $a_n \models q_n$ and $d^{\boldsymbol{M}}(a_n, a_{n+1}) \leq \frac{1}{2^n}$. The fact that $(M_x, d^{\boldsymbol{M}})$ is a complete metric space (Remark 1.26), yields some $a \in M_x$ such that $a_n \to a$. Hence $q_n \to \mathrm{tp}^{\boldsymbol{M}}(a)$, which in turn implies $p_n \to \mathrm{tp}^{\boldsymbol{M}}(a)$. □

Remark 1.81. *The space* $\mathrm{S}_x(T)$ *with the metric topology can be also viewed as the quotient* $M_x/\{(a,b) : \exists h \in \mathrm{Aut}(\boldsymbol{M}), h(a) = b\}$, *for* $\boldsymbol{M}$ *an* $\max(\aleph_0, |x|^+)$*-monster model, with the distance between two classes* p, q *given by* $d(p,q) = \inf\{d^{\boldsymbol{M}}(a,b) : a \in p, b \in q\}$.

Remark 1.82. *Let* p, q *be two* x*-types over a parameter set* $C \subseteq M$, *and* y *be a multi-variable of the same type as* x. *Let* $r = \sup\{d(p|_F, q|_F) : F \subseteq C \text{ finite}\}$. *We observe that for every finite* $F \subseteq C$ *we have* $d(p|_F, q|_F) \leq d(p,q)$, *from which it follows* $r \leq d(p,q)$. *On the other hand, by the definition of* r *we have that*

$$p(x) \cup q(y) \cup \{d(x,y) \leq r\}$$

is finitely satisfiable, hence it is a type, implying $d(p,q) \leq r$. *Summarizing, we obtained*

$$d(p,q) = \sup\{d(p|_F, q|_F) : F \subseteq C \text{ finite}\}.$$

Since formulas satisfy a modulus of continuity, for any $\mathcal{L}$-formula φ the function $\dot{\varphi} : \mathrm{S}_x(T) \to [0,1]$ is uniformly continuous with respect to the metric topology. Furthermore, by Proposition 1.61, every continuos function $f : \mathrm{S}_x(T) \to [0,1]$, with respect to the logic topology, is uniformly continuous with respect to the metric topology. We end this introduction to type spaces with the following proposition:

Proposition 1.83. *Let $A \subseteq B \subseteq M$. Then the map $\pi : \mathrm{S}_x(B) \to \mathrm{S}_x(A)$ given by*

$$p \mapsto p|_A = \{E(x) : E(x) \in p, E(x) \text{ is an } \mathcal{L}_A(x)\text{-condition}\}$$

satisfies:

(i) *π is continuous and surjective with respect to the logic topologies;*

(ii) *π is 1-Lipschitz continuous, with respect to the metric topologies;*

(iii) *if A is dense in B then π is a homeomorphism with respect to the logic topology and an isometry with respect to the metric topologies.*

Item (*i*) and (*ii*) are straighforward while for (*iii*) we use Proposition 1.12 to prove that every $p \in \mathrm{S}_x(A)$ extends uniquely to a type in $\mathrm{S}_x(B)$.

The logic mechanism on type spaces

Throughout this subsection let T be a theory (not necessarily complete), $\boldsymbol{M}$ a model of T, $x = (x_i)_{i\in I}$ a multivariable, and let (y_n) be a sequence of distinct single variables. Our aim is to present in a rigorous way how one can also view formulas and terms defined on the type spaces.

Definition 1.84. Let $t = (t_i)_{i\in I}$ be a sequence of $\mathcal{L}$-terms, and $p(x)$ be a type in $\mathrm{S}_x(T)$. Then we define $p(t)$ to be the set of all conditions which are obtained by the substitution of t_i in place of x_i, for each i, in the conditions of $p(x)$. Similarly, for an $\mathcal{L}(x)$-formula φ, we define $\varphi(t)$ the formula obtained by the substitution of t_i for x_i for each i.

For an $\mathcal{L}(x)$-term t, we define $\dot{t}_x : \mathrm{S}_x(T) \to \mathrm{S}_{y_1}(T)$, also abbreviated as $\dot{t}$, by $p \mapsto q$ such that $q(t) \subseteq p$. This map is well-defined and continuous, because for any $\mathcal{L}(x)$-formula φ we have $(\dot{t}_x)^{-1}[\varphi(y) < \varepsilon] = [\varphi(t) < \varepsilon]$. For $\mathcal{L}(x)$-terms $t_1, \dots, t_n$ we define the map $(t_1, \overset{\bullet}{\dots}, t_n)_x : \mathrm{S}_x(T) \to \mathrm{S}_{y_1\dots y_n}(T)$ by

$$p \mapsto q \text{ such that } q(t_1, \dots, t_n) \subseteq p.$$

This is also well-defined and continuous. Clearly these functions depend on the multivariable x. If y is a multivariable containing x, then it is easy to see that $(t_1, \overset{\bullet}{\dots}, t_n)_y = (t_1, \overset{\bullet}{\dots}, t_n)_x \circ \eta_{y,x}$ (notation from Corollary 1.64). If t is a constant symbol c and x is the empty variable then

$$\mathrm{S}_\emptyset(T) = \{p : p \text{ is a complete theory with } T \subseteq p\}$$

and by definition we obtain the map $\dot{c} : \mathrm{S}_\emptyset(T) \to \mathrm{S}_{y_1}(T)$ which sends $p \in \mathrm{S}_\emptyset(T)$ to the unique y_1-type $q(y_1)$ such that $q(c) \subseteq p$. We also have the following:

Proposition 1.85. *Let $t_1, \dots, t_n$ be $\mathcal{L}(x)$-terms and f be an n-ary function symbol. Then we have*

$$(f(t_1, \overset{\bullet}{\dots}, t_n))_x = (fy_1 \overset{\bullet}{\dots} y_n)_{y_1,\dots,y_n} \circ (t_1, \overset{\bullet}{\dots}, t_n)_x.$$

$$\begin{array}{ccc} M_x & \xrightarrow{t^{M}} & M_{y_1} \\ {\scriptstyle \iota_x}\downarrow & & \downarrow{\scriptstyle \iota_{y_1}} \\ \mathrm{S}_x(T) & \xrightarrow{\dot{t}_x} & \mathrm{S}_{y_1}(T) \end{array} \qquad\qquad \begin{array}{ccc} M_x & \xrightarrow{\varphi^{M}} & [0,1] \\ {\scriptstyle \iota_x}\downarrow & & \downarrow{\scriptstyle id} \\ \mathrm{S}_x(T) & \xrightarrow{\dot{\varphi}_x} & [0,1] \end{array}$$

Diagram 1 Diagram 2

The most important property is that for any $\mathcal{L}(x)$-term t, Diagram 1 commutes. Similarly, for any $\mathcal{L}(x)$-formula $\varphi(x)$, Diagram 2 also commutes.

Remark 1.86. *It is not necessary to restrict ourselves to only the sequence of single variables* (y_n)*, we can choose any other sequence of single variables* (z_n)*, and similarly define the map* $\dot{t}_x$ *but with co-domain* $\mathrm{S}_{z_1,\ldots,z_n}(T)$*. The notation* $\dot{t}_x$ *has no mention of the respective sequence of single variables chosen, even though this is not exact, this notation is preferred, as we just have to choose some sequences of single variables in order for some compositions of functions of the type* $\dot{t}$ *to make sense.*

We now arrive at the main proposition of this section, whose proof follows directly from the definitions.

Proposition 1.87. *Let* $\varphi(y_1, \ldots, y_n)$ *be an* $\mathcal{L}$*-formula,* $t_1, \ldots, t_n$ *be* $\mathcal{L}(x)$*-terms, and* $\psi = \varphi(t_1, \ldots, t_n)$ *(each* y_m *is substituted with* t_m*). Then*

$$\dot{\psi}_x = \dot{\varphi}_{y_1 \ldots y_n} \circ (t_1, \dot{\ldots}, t_n)_x$$

Example 1.88. *Proposition 1.87 helps us understand how to break down the function* $\dot{\varphi}$ *into its "buildings blocks". For example, suppose* x, y *are single variables, and let* $\varphi = d(f(x), y)$*, where* f *is an unary function symbol of* $\mathcal{L}$*. The free variables of* φ *are* x *and* y*, hence* $\dot{\varphi} : \mathrm{S}_{xy}(T) \to [0,1]$*. But it breaks down as, using Remark 1.86,* $(f\dot{x}, y) : \mathrm{S}_{xy}(T) \to \mathrm{S}_{xy}(T)$ *which maps* $p(x,y)$ *into the unique complete* xy*-type* $q(x,y)$ *such that* $q[f(x), y] \subseteq p$ *(the substitution of* x *with* $f(x)$ *for each condition in* $q(x,y)$*), and composing this with* $d(\dot{x}, y) : \mathrm{S}_{xy}(T) \to [0,1]$ *gives us* $\dot{\varphi}$*.*

We end this subsection with the following theorem, which has an immediate proof.

Theorem 1.89. *Let* φ *be a closed formula. Then*

$$T \models \varphi = 0 \Longleftrightarrow \dot{\varphi} : \mathrm{S}_\emptyset(T) \to [0,1] \text{ is the constant 0 map.}$$

2 Model Theory

In this chapter we present the definitions and basic properties of definable predicates, definable sets, definable functions, definable elements, algebraic elements and some additional topics such as back-and-forth systems and a quantifier elimination criteria. Our approach is based on the framework of type spaces, offering a slightly different perspective and presentation from that of [BBHU08].

2.1 Definability

Throughout this subsection let T be a complete $\mathcal{L}$-theory, $\boldsymbol{M}$ a model of T, and x a multivariable.

Before introducing definability concepts in continuous logic, we first establish some preliminary observations and definitions. The logic topology is assumed to be the default topology on $\mathrm{S}_x(T)$. On M_x, unless otherwise specified, we consider the topology induced by the metric $d^{\boldsymbol{M}}$. Then, by Proposition 1.12, the map $\iota_x = \iota_x^{\boldsymbol{M}} : M_x \to \mathrm{S}_x(T)$ defined by $\iota_x(a) = \mathrm{tp}(a)$ is continuous, and $\iota_x(M_x)$ is dense in $\mathrm{S}_x(T)$. Therefore, for any function $f : M_x \to [0,1]$, exists at most one continuous function $F : \mathrm{S}_x(T) \to [0,1]$ such that $f = F \circ \iota$; if such an F exists, then f is continuous.

Definition 2.1. An *x-predicate* (or simply a *predicate*) is a function $P : M_x \to [0,1]$.

For any relation symbol R, the intepretation $R^{\boldsymbol{M}}$ is a predicate. For any $\mathcal{L}(x)$-formula φ, $\varphi^{\boldsymbol{M}} = \varphi_x^{\boldsymbol{M}}$ is a predicate. Observe that $\varphi^{\boldsymbol{M}} = \dot{\varphi} \circ \iota_x$. Inspired by this observation we define:

Definition 2.2. We say $P : M_x \to [0,1]$ is a 0-*definable x-predicate* on $\boldsymbol{M}$, or simply a 0-*definable predicate* on $\boldsymbol{M}$, if there exists a continuous function $\Phi = \Phi_P : \mathrm{S}_x(T) \to [0,1]$ such that $P = \Phi \circ \iota_x$. For a parameter set $A \subseteq M$, we call $P : M_x \to [0,1]$ an *A-definable x-predicate* on $\boldsymbol{M}$ if P is a 0-definable x-predicate on $\boldsymbol{M}_A$. Finally, we say that $P : M_x \to [0,1]$ is a *definable x-predicate* on $\boldsymbol{M}$ if it is an M-definable x-predicate on $\boldsymbol{M}$.

Let $\boldsymbol{M} \preccurlyeq \boldsymbol{N}$, P be a 0-definable x-predicate on $\boldsymbol{M}$, and $\Phi : \mathrm{S}_x(T) \to [0,1]$ such that $P = \Phi \circ \iota_x^{\boldsymbol{M}}$. Then we obtain $Q = \Phi \circ \iota_x^{\boldsymbol{N}}$ is a 0-definable x-predicate on $\boldsymbol{N}$, with the property that $Q \restriction_{M_x} = P$. Because $\iota_x^{\boldsymbol{M}}[M_x]$ is dense in $\mathrm{S}_x(T)$, it follows that Q is the unique 0-definable x-predicate on $\boldsymbol{N}$ such that $Q \restriction_{M_x} = P$. Therefore, the 0-definable x-predicates of $\boldsymbol{M}$ are in an unique correspondence with the 0-definable x-predicates of $\boldsymbol{N}$.

By Proposition 1.61, Φ is the uniform limit of a sequence of functions $\dot{\varphi}_n$, where φ_n is an $\mathcal{L}(x)$-formula. Since each φ_n has a finite set of free variables, it follows that there exists a countable multivariable $y \subseteq x$ such that Φ factors through $\mathrm{S}_y(T)$. Therefore we obtain:

Proposition 2.3. *For any 0-definable x-predicate P there exists a countable multivariable $y \subseteq x$ and a 0-definable y-predicate Q such that $P = Q \circ \pi_{x,y}$, where $\pi_{x,y} : M_x \to M_y$ is the projection map.*

Proof. Let $\Phi_x : \mathrm{S}_x(T) \to [0,1]$ be the continuous function such that $P = \Phi_x \circ \iota_x$. Then there exists a countable multivariable $y \subseteq x$ and $\Phi_y : \mathrm{S}_y(T) \to [0,1]$ continuous such that $\Phi_x = \Phi_y \circ \eta_{x,y}$. Then, $Q = \Phi_y \circ \iota_y$ is a 0-definable y-predicate such that $P = Q \circ \pi_{x,y}$. $\square$

As a consequence, the 0-definable predicates are uniquely determined by a countable number of variables, making it reasonable to restrict x to be countable in the definition of a 0-definable predicate. However, for now, we will not impose any specific cardinality constraints on x.

Definition 2.4. Let X be any set and Y a metric space. A sequence of functions $(f_n) \subseteq Y^X$ is called *cauchy*, if it is a cauchy sequence in the metric space Y^X endowed with the supremum metric. Similarly, a sequence of $\mathcal{L}(x)$-formulas (φ_n) is called *cauchy* if $(\dot{\varphi}_n)$ is cauchy.

Remark 2.5. *We can also, equivalently, define a sequence of $\mathcal{L}(x)$-formulas (φ_n) to be cauchy if for all $\varepsilon > 0$ exists a natural number $k > 0$ such that for all $n, m \geq k$ we have $\sup_x |\varphi_n - \varphi_m| \leq \varepsilon$ belongs to T. By Theorem 1.89, this is equivalent to the sequence $\dot{\varphi}_n$ being cauchy.*

With notation as above if Φ is the uniform limit of $(\dot{\varphi}_n)$ then $(\dot{\varphi}_n)$ is a cauchy sequence, and is straightfoward to see that $(\varphi_n^{\boldsymbol{M}})$ is also cauchy and $P = \lim \varphi_n^{\boldsymbol{M}}$. The property that $(\varphi_n^{\boldsymbol{M}})$ is a cauchy sequence is described by the theory T (for example $\sup_x |\varphi_n - \varphi_m| \leq \varepsilon$), therefore for $\boldsymbol{M} \preccurlyeq \boldsymbol{N}$, $(\varphi_n^{\boldsymbol{N}})$ is also a cauchy sequence. Since $\varphi_n^{\boldsymbol{N}} = \dot{\varphi}_n \circ \iota_x^{\boldsymbol{N}}$ it is clear that $Q = \lim \varphi_n^{\boldsymbol{N}}$ is the corresponding 0-definable predicate of P in $\boldsymbol{N}$. Therefore, given T (forgetting about the model $\boldsymbol{M}$), we can also define a 0-definable predicate by a sequence of $\mathcal{L}(x)$-formulas (φ_n) such that for all $\varepsilon > 0$ exists a natural number $k > 0$ such that for all $n, m \geq k$ we have $\sup_x |\varphi_n - \varphi_m| \leq \varepsilon$ belongs to T. Unfortunately this will not be an unique assignment (from a sequence of "cauchy" formulas to a 0-definable predicate), as more sequences of formulas can give the same 0-definable predicate.

Remark 2.6. *Since 0-definable predicates are uniform limits of cauchy sequences of formulas, each of which has a modulus of continuity, it follows that 0-definable predicates also have a modulus of continuity. Therefore, when extending a language by adding a relation symbol whose interpretation in a model is a 0-definable predicate, then we also need to specify to the language a modulus of continuity that the interpretation is required to satisfy.*

Proposition 2.7. *Suppose x is finite, $\boldsymbol{M} \preccurlyeq \boldsymbol{N}$. Let P be a 0-definable x-predicate on $\boldsymbol{M}$ and Q be an uniformly continuous predicate on N_x. Let R be a distinct relation symbol from $\mathcal{L}$, for which we interpret $R^{\boldsymbol{M}} = P$. Then $(\boldsymbol{M}, P) \preccurlyeq (\boldsymbol{N}, Q)$ (here $Q = R^{\boldsymbol{N}}$) if and only if Q is the 0-definable predicate corresponding to P.*

Proof. Suppose $(\boldsymbol{M}, P) \preccurlyeq (\boldsymbol{N}, Q)$ and let $(\varphi_n(x))$ be a sequence of formulas such that P is the uniform limit of $(\varphi_n^{\boldsymbol{M}})$. Then, by $\boldsymbol{M} \preccurlyeq \boldsymbol{N}$, Q is the uniform limit of $(\varphi_n^{\boldsymbol{N}})$, immediately implying Q is the 0-definable predicate corresponding to P. For the reverse direction, we use a sequence (φ_n) defining the predicate P in order to prove by induction that for any $(\mathcal{L} \cup \{R\})(y)$-formula ψ (y any multivariable) and any $\varepsilon > 0$ exists an $\mathcal{L}(y)$-formula φ such that

$$|\psi^{\boldsymbol{M}}(a) - \varphi^{\boldsymbol{M}}(a)| \leq \varepsilon \text{ and } |\psi^{\boldsymbol{N}}(b) - \varphi^{\boldsymbol{N}}(b)| \leq \varepsilon,$$

for any $a \in \boldsymbol{M}_y$ and $b \in \boldsymbol{N}_y$. By $\varepsilon \to 0$, we obtain that $\psi^{\boldsymbol{M}}(a) = \psi^{\boldsymbol{N}}(a)$ for any $a \in \boldsymbol{M}_y$. □

This result easily generalizes to:

Proposition 2.8. *Let $\boldsymbol{M} \preccurlyeq \boldsymbol{N}$, and $(P_j)_{j \in J}$ be a collection of finite 0-definable predicates on $\boldsymbol{M}$. Let $(R_j)_{j \in J}$ be a collection of disjoint relation symbols, for which we interpret for each j, $R_j^{\boldsymbol{M}} = P_j$. Then $(\boldsymbol{M}, (P_j)_{j \in J}) \preccurlyeq (\boldsymbol{N}, (Q_j)_{j \in J})$ (here $Q_j = R_j^{\boldsymbol{N}}$) if and only if, for each j, Q_j is the 0-definable predicate corresponding to P_j.*

The next proposition studies the connection between 0-definable predicates and continuous connectives.

Proposition 2.9. *Let γ be an n-ary connective, $P, P_1, \ldots, P_n$ be 0-definable x-predicates, and y a quantifiable multivariable with $y \subseteq x$. Then we have the following:*

(i) $\gamma(P_1, \ldots, P_n)$ *is a 0-definable x-predicate;*

(ii) $\inf_y P : M_{x \setminus y} \to [0,1]$ *defined by* $\inf_y P(b) := \inf\{P(a) : b = a|_{x \setminus y},\ a \in M_x\}$, *or simply by*

$$\inf_y P(b) = \inf_{b = a|_{x \setminus y}} P(a),$$

is a 0-definable predicate.

Proof. The predicate $\gamma(P_1, \ldots, P_n)$ obviously corresponds to the continuous function $\gamma \circ (\Phi_{P_1}, \ldots, \Phi_{P_n})$, hence it is a 0-definable x-predicate. For item (ii), let (φ_n) be a sequence of $\mathcal{L}(x)$-formulas such that $(\varphi_n^{\boldsymbol{M}})$ uniformly converges to P. For each n, we can consider all the free variables of φ_n to be quantifiable. Then all it remains is to observe that $(\inf_y \varphi_n)$ (Remark 1.66) converges uniformly to $\inf_y P$. (Here we use notation from Remark 1.9). □

Similar to Proposition 1.65, we have the following:

Proposition 2.10. *Let $\Phi : \mathrm{S}_x(T) \to [0,1]$ be a continuous function corresponding to the 0-definable x-predicate P, and a quantifiable multivariable $y \subseteq x$. Then we have that*

$$\inf_y \Phi : \mathrm{S}_{x\setminus y}(T) \to [0,1] \text{ defined by } \inf_y \Phi(q) := \{\Phi(p) : p|_{x\setminus y} = q, p \in \mathrm{S}_x(T)\},$$

is a continuous function corresponding to the 0-definable $x \setminus y$-predicate $\inf_y P$.

A proof of this follows by applying Remark 1.66, to a sequence of $\mathcal{L}(x)$-formulas (φ_n) such that $(\dot{\varphi_n})$ converges uniformly to Φ.

Suppose $x = (x_n)_n$ is a countable quantifiable multivariable and let $\Phi : \mathrm{S}_x(T) \to [0,1]$ be continuous. Let $P : M_x \to [0,1]$ be the 0-definable x-predicate corresponding to Φ. Then by using the fact that P is the limit of $(\varphi_n^{\boldsymbol{M}})$ for some cauchy sequence (φ_n) of $\mathcal{L}$-formulas, we observe that for all $\varepsilon > 0$ exists m such that for $a, b \in M_x$ if $a_i = b_i$ for all $0 \leq i \leq m$ then $|P(a) - P(b)| \leq \varepsilon$. As a simple consequence, we obtain

$$P(a) = \lim_n \left(\left(\inf_{(x_m)_{m>n}} P \right) (a_0, \ldots, a_n) \right)$$

for all $a \in M_x$. If P, Q are two different 0-definable x-predicates, then we observe that exists $y \subseteq x$, with $x \setminus y$ a finite multivariable, such that $\inf_y P$ is different from $\inf_y Q$. Therefore the assignment from P a 0-definable x-predicate to a sequence of finitary 0-definable predicates $(\inf_y P)_y$, where y ranges over $y \subseteq x$ such that $x \setminus y$ is finite, is injective. This assignment clearly takes into consideration the model $\boldsymbol{M}$. But, we also have, by the same reasons, the following injective assigment from $\Phi : \mathrm{S}_x(T) \to [0,1]$ to a sequence of "finitary" continuous functions $(\inf_y \Phi)_y$ where y ranges over $y \subseteq x$ such that $x \setminus y$ is finite. Therefore combining these facts with Proposition 2.7, we obtain:

Proposition 2.11. *Suppose $x = (x_n)_n$ is a countable quantifiable multivariable, $\boldsymbol{M} \preccurlyeq \boldsymbol{N}$, and P a 0-definable x-predicate on $\boldsymbol{M}$. Then $(\boldsymbol{M}, (\inf_y P)_y) \preccurlyeq (\boldsymbol{N}, (Q_y)_y)$, where y ranges over $y \subseteq x$ with $x \setminus y$ finite, if and only $Q_y = \inf_y Q$ where Q is the 0-definable predicate on $\boldsymbol{N}$ corresponding to P.*

Remark 2.12. *One interesting fact is that 0-definable predicates, unlike formulas, are also closed under infinitary connectives. For example if $\gamma : [0,1]^J \to [0,1]$ is continuous and P_n is a 0-definable x-predicate, for each n, then $\gamma \circ (P_j)_{j \in J}$ is a 0-definable x-predicate since it corresponds to the continuous function $\gamma \circ (\Phi_{P_j})_{j \in J} : \mathrm{S}_x(T) \to [0,1]$. We thus encounter a remarkable phenomenon: on one hand we have formulas involving only finite connectives and finite quantifiers, a "finitary logic", and on the other side we have the 0-definable predicates on which we can apply infinitary connectives and infinitary (countable) quantifiers, hence an "infinitary logic". Clearly, this infinitary logic has the property that it can be "approximated" as well as we want by the finitary logic.*

For future reference, we state the following proposition, which follows easily from the definitions.

Proposition 2.13. *Suppose x is finite.*

(i) *For any 0-definable x-predicate P and $(\boldsymbol{M}, P) \preccurlyeq (\boldsymbol{N}, Q)$ we have that* $\mathrm{im}(P)$ *is a dense subset of* $\mathrm{im}(Q)$*, and* $\mathrm{im}(P)$ *is a dense subset of* $\mathrm{im}(\Phi_P)$*;*

(ii) *Let P, Q be 0-definable x-predicate with the same zero-sets, i.e. $P^{-1}(0) = Q^{-1}(0)$, then P and Q satisfy the statements*

$$\forall \varepsilon > 0, \exists \delta > 0, \forall a \in M_x,\ P(a) \leq \delta \implies Q(a) \leq \varepsilon, \ \textit{and}$$
$$\forall \varepsilon > 0, \exists \delta > 0, \forall a \in M_x,\ Q(a) \leq \delta \implies P(a) \leq \varepsilon.$$

Proof. For item (ii), $P^{-1}(0) = Q^{-1}(0)$ implies that $P(x) = 0 \models Q(x) = 0$. By applying Corollary 1.44 (ii), we conclude that P and Q satisfy both statements. □

Before presenting the continuous version of Svenonius Theorem, we first establish a topological lemma. Let X be a set equipped with two topologies, τ_1 and τ_2, where τ_1 is hausdorff, τ_2 is compact, and $\tau_1 \subseteq \tau_2$. Then as a consequence of the fact that the identity function $\mathrm{id} : (X, \tau_2) \to (X, \tau_1)$ is a continuous bijection from a compact space to a hausdorff space, we obtain $\tau_1 = \tau_2$ [Roy10, Proposition 19, p.235]. Thus, in a hausdorff topological space, any finer compact topology must coincide with it.

Now, let $f : X \to Y$ (Y any hausdorff space) be a continuous function. Define the quotient set $X/\{f\} = X/\{a \sim b : a, b \in X, f(a) = f(b)\}$ with the natural projection $\pi : X \to X/\{f\}$. Then, there exists an unique map $\tilde{f} : X/\{f\} \to Y$ such that $f = \tilde{f} \circ \pi$, and as a consequnce of the previous observation we have that the quotient topology on $X/\{f\}$ is equivalent to the initial topology generated by $\tilde{f}$. The space $X/\{f\}$ is compact, hausdorff, and the function $\tilde{f}$ is a homeomorphism on its image.

Lemma 2.14. *Let X be a compact hausdorff space, Y and Z be two hausdorff spaces, and $\tau : X \to Z$ be a continuous surjection. Suppose $f : X \to Y$ is a continuous function for which exists $\tilde{f} : Z \to Y$ such that $f = \tilde{f} \circ \tau$. Then $\tilde{f}$ is continuous.*

Proof. All we have to observe is that as a consequence of $f = \tilde{f} \circ \tau$ and the previous observations, the topological space $X/\{f\}$ is a quotient space of $X/\{\tau\}$, therefore obtaining that $\tilde{f}$ is continuous. □

Definition 2.15. Using the notation from Lemma 2.14, we also denote $\tilde{f}$ as f_Z.

Theorem 2.16 (Svenonius Theorem)**.** *Suppose x is finite, R is a relation symbol distinct from $\mathcal{L}$, and let $R^{\boldsymbol{M}} = P : M_x \to [0,1]$ be an uniformly continuous predicate. Then P is a 0-definable predicate in $\boldsymbol{M}$ if and only if for all $(\boldsymbol{N}, Q) \succcurlyeq (\boldsymbol{M}, P)$ the predicate Q is invariant under all automorphisms of $\boldsymbol{N}$.*

This theorem is stated and proved in [BBHU08, Corollary 9.11], where saturation is used to prove the continuity of a function. Here we provide a topological argument of this fact.

Proof. If P is a 0-definable predicate in $\boldsymbol{M}$ then it is a limit of formulas, and since formulas are invariant under automorphisms, it follows that Q is also invariant under all automorphisms of $\boldsymbol{N}$. For the converse direction let $T_R = \mathrm{Th}(\boldsymbol{M}, P)$, and consider the function $\dot{R} : \mathrm{S}_x(T_R) \to [0,1]$. Let $p, q \in \mathrm{S}_x(T_R)$ with $p|_{\mathcal{L}} = q|_{\mathcal{L}}$. By taking an elementary extension $(\boldsymbol{N}, Q)$ of $(\boldsymbol{M}, P)$, which is $\aleph_0$-saturated and $\aleph_0$-strongly homogeneous in every reduct of $\mathcal{L} \cup \{R\}$, we obtain that for $a, b \in N_x$ such that $a \models p, b \models q$ there exists an automorphism of $\boldsymbol{N}$ sending a to b. Therefore $\dot{R}(p) = \dot{R}(q)$, and as a consequence $\dot{R}$ factors through the continuous surjection $\pi : \mathrm{S}_x(T_R) \to \mathrm{S}_x(T)$ which sends $p \to p|_{\mathcal{L}}$. By Lemma 2.14 there exists a continuous $\tilde{R} : \mathrm{S}_x(T) \to [0,1]$ such that $\dot{R} = \tilde{R} \circ \pi$. This implies that $P = \tilde{R} \circ \iota_x^{\boldsymbol{M}}$, concluding P is a 0-definable predicate on M_x. □

More on definable predicates

Throughout this subsection x is a finite multivariable. Our next aim is to prove that 0-definable predicates have the exact same behavior as formulas, and that they can be viewed as syntactic objects as well.

Let (φ_n) be a Cauchy sequence of $\mathcal{L}(x)$-formulas, and P the corresponding 0-definable predicate on $\boldsymbol{M}$. Consider a type $p \in \mathrm{S}_x(T)$. Then p already knows the "truth value of P" : the value of $\Phi_P(p)$ is the limit of $\dot{\varphi}_n(p)$. Therefore, in a similar fashion as with formulas, we also denote $\dot{P} := \Phi$ and the truth value of P in p by $P^p = \dot{P}(p)$. Clearly, for an n-ary connective γ, and a sequence $P_1, \ldots, P_n$ of 0-definable x-predicates, by Proposition 2.9, $\gamma(\dot{P}_1, \ldots, \dot{P}_n) = (\gamma(P_1, \ldots, P_n))^{\boldsymbol{\cdot}}$, and for a multivariable $y \subseteq x$ we have $\inf_y \dot{P} = (\inf_y P)^{\boldsymbol{\cdot}}$. Even though we already know that $\inf_y \dot{P} = (\inf_y P)^{\boldsymbol{\cdot}}$, in order to show the similarity of predicates with formulas, starting from the assumptions that $\inf_y P$ is a 0-definable predicate, being the limit of the cauchy sequence of formulas $\inf_y \varphi_n$, and that $\inf_y \dot{P} : \mathrm{S}_x(T) \to [0,1]$ is continuous, we will prove $\inf_y \dot{P} = (\inf_y P)^{\boldsymbol{\cdot}}$. Hence, using the same notations:

Proposition 2.17. *We have that* $(\inf_y P)^{\boldsymbol{\cdot}} = \inf_y \dot{P}$ *and for each* $q \in \mathrm{S}_{x\setminus y}(T)$ *there exists* p *containing* q *such that* $(\inf_y P)^{\boldsymbol{\cdot}}(q) = \dot{P}(p)$.

Proof. This is almost the same proof as for formulas from (ii) of Proposition 1.65. Let $q \in \mathrm{S}_{x\setminus y}(T)$ and p be any x-type extending q. Then $(\inf_y P)^q$ is the limit of $(\inf_y \varphi_n)^q$. Since $q \subseteq p$, $(\inf_y \varphi_n)^q = (\inf_y \varphi_n)^p$, and then obviously $(\inf_y \varphi_n)^p \leq \varphi_n^p$. We obtain

$$(\inf_y P)^{\boldsymbol{\cdot}}(q) = \lim_n (\inf_y \varphi_n)^q = \lim_n (\inf_y \varphi_n)^p \leq \lim_n \varphi_n^p = \dot{P}(p).$$

Hence, $(\inf_y P)^{\boldsymbol{\cdot}}(q) \leq \inf\{\dot{P}(p) : q \subseteq p, p \in \mathrm{S}_x(T)\} = \inf_y \dot{P}(q)$. Let $r = (\inf_y P)^{\boldsymbol{\cdot}}(q)$. For each n we can consider $\sup |P - \varphi_n| \leq \frac{1}{2^n}$. Then, by using a realization of q in a model of T, we observe that the set of conditions $q \cup \{|\varphi_n(x) - r| \leq \frac{1}{2^{n-1}} : n\}$ is finitely satisfiable, therefore it can be extended to a type p. Clearly, we have that $\dot{P}(p) = r$. □

Remark 2.18. *For each* $a \in M_x$ *we have* $P(a) = r$ *if and only if* $\dot{P}(\mathrm{tp}^{\boldsymbol{M}}(a)) = r$. *Which, by adding a relation symbol for* P *we could write*

$$\boldsymbol{M} \models P(a) = r \Leftrightarrow \dot{P}(\mathrm{tp}^{\boldsymbol{M}}(a)) = r.$$

Inspired by the above remark, for each n, let S_n^r be a set of n-ary relation symbols disjoint from $\mathcal{L}^r$, which is in bijection with $\mathrm{C}(\mathrm{S}_n(T),[0,1])$. For a relation symbol $P \in S_n^r$ we denote by $\dot{P}$ the corresponding element in $\mathrm{C}(\mathrm{S}_n(T),[0,1])$. Let $\mathcal{L}^m$ be the language $\mathcal{L}$ to which we adjoin the relation symbols S_n^r for each n, and the corresponding modulus of continuities. We define T^m to be the completion (the unicity is mentioned in Theorem 2.19) of

$$\Lambda = T \cup \{\sup_x |P-\varphi| \leq \varepsilon : n, P \in S_n^r, \varepsilon > 0,\ x \text{ a multivariable with } |x| = n,$$
$$\varphi \in \mathcal{L}(x),\ \sup_{p \in \mathrm{S}_x(T)} |\dot{P} - \dot{\varphi}| \leq \varepsilon\},$$

where in the obvious way we identify $\mathrm{S}_x(T)$ with $\mathrm{S}_n(T)$.

It is easy to see that $\boldsymbol{M}$ can be uniquely extended to an $\mathcal{L}^m$-structure satisfying T^m. Moreover, for each $P \in S_n^r$, we have that $\dot{P^{\boldsymbol{M}}} = \dot{P}$. The next proposition presents the connections between T and T^m.

Theorem 2.19. *The following hold:*

(i) Λ has an unique expansion to a complete theory of $\mathcal{L}^m$;

(ii) Every type $p \in \mathrm{S}_x(T)$ can be uniquely extended to a type in $\mathrm{S}_x(T^m)$;

(iii) the correspondence $\mathrm{S}_x(T) \to \mathrm{S}_x(T^m)$ from item (ii) is a homeomorphism;

(iv) each continuous function $\mathrm{S}_x(T^m) \to [0,1]$ is $\dot{R}$ for some relation symbol of $\mathcal{L}^m$.

We close this subsection with the following remark:

Remark 2.20. *The big advantage of adding the language $\mathcal{L}^m$ and $\mathcal{T}^m$ is that we can freely use compactness theorem on conditions involving 0-definable predicates, without mentioning the use of a cauchy sequences of formulas defining the 0-definable predicates. For example: Let P be a 0-definable x-predicate of $\boldsymbol{M}$ and $a \in M_x$. We will abuse notation for P as a relation symbol and as a 0-definable predicate. Suppose also that $\inf_x P = r$, for some $r \in [0,1]$. Then $\{|P(x) - r| \leq \frac{1}{n} : n > 0\}$ is finitely satisfiable in the structure $(\boldsymbol{M}, P)$, hence there exists an elementary extension $(\boldsymbol{M}, P) \preccurlyeq (\boldsymbol{N}, P)$ for which exists some $a \in N_x$ such that $P(a) = r$. Using a similar argument we can provide a similar style proof of Proposition 2.17.*

Definable sets in continuous logic

Throughout this subsection x is a finite multivariable.
In the search for a definition of the concept of a definable set, we should follow a few principles:

(i) just as in the case with definable predicates, where on the one hand there is the definable predicate on the structure and the unique corresponding continuous function from the type space to $[0,1]$, a definable set should also have a unique corresponding object involving the type space;

(ii) there should be a bijective correspondence between the definable sets between two models of T, one being an elementary extension of the other;

(iii) just as in the case with definable predicates, where there is a set of formulas that can verify if something is a definable predicate, there should also be a set of formulas which verify that something is a definable set;

(iv) the definition of the concept of a definable set should involve the logic topology in a crucial way, as we want to be able to uniformly approximate the concept with formulas.

In classical logic, given an element a we can know if an element is in the definable set or not by checking the truth value of a formula $\varphi(a)$, but as we are working in continuous logic where truth values can be anything from $[0,1]$, given an element it makes more sense to know how far it is from the definable set. Therefore:

Definition 2.21. We call a subset K of $\mathrm{S}_x(T)$ *0-definable* if the function

$$p \mapsto d(p,K)\colon S_x(T) \to [0,1]$$

is continuous, where d is as in Definition 1.79.

Definition 2.22. We call a subset A of M_x *0-definable* if $A = \iota_x^{-1}(K)$ for K a 0-definable subset of $\mathrm{S}_x(T)$.

Before continuing, we need some expressions or formulas which can verify that a continuous function is of the type $d(x,Y)$ (Y is any subset of some metric space), such functions we call *distance functions*.

Proposition 2.23. *Let X be a complete metric space with diameter ≤ 1. Then a continuous function $f : X \to [0,1]$ is a distance function if and only if either $f = 1$ or the following two expressions hold:*

(E1) $\sup_{x\in X} \inf_{y\in X} \max(f(y), |f(x) - d(x,y)|) = 0$*;*

(E2) $\sup_{x\in X} |f(x) - \inf_{y\in X} \min(f(y) + d(x,y))| = 0.$

Proof. Suppose f is a distance function with $K = f^{-1}(0)$. If $K = \emptyset$ then $f = 1$ if $K \neq \emptyset$ then $E1, E2$ hold.

For the reverse direction suppose $E1, E2$ hold and let $K = f^{-1}(0)$. By $E2$ we have that $f(x) \leq f(y) + d(x,y)$, and by letting y ranging from K, we obtain $f(x) \leq d(x,K)$. Now, let $x \in X$ and $\varepsilon > 0$. Applying $E1$ repeatedly we obtain a sequence (x_n) with $x_0 = x$ such that for each $n > 0$ we have $f(x_n) \leq \frac{\varepsilon}{2^n}$ and for each $n \geq 0$ we have $|f(x_n) - d(x_n, x_{n+1})| \leq \frac{\varepsilon}{2^n}$. Then (x_n) is cauchy. Let y be its limit. By using the continuity of f, the triangle inequality for d and $|\cdot|$, we obtain that $f(y) = 0$, $|f(x) - d(x,x_1)| \leq \frac{\varepsilon}{2}$ and $d(x_1,y) \leq 2\varepsilon$. Hence $d(x,K) \leq d(x,y) \leq f(x) + 3\varepsilon$, letting $\varepsilon \to 0$, we obtain $f(x) = d(x,K)$. $\square$

Remark 2.24. *If $K \subseteq \mathrm{S}_x(T)$ is 0-definable, then there is a unique distance function Φ on $\mathrm{S}_x(T)$ with the metric topology, such that $K = \Phi^{-1}(0)$. Moreover Φ is either equal to 1 or satisfies*

(i) $\sup_{p\in \mathrm{S}_x(T)} \inf_{q\in \mathrm{S}_x(T)} \max(\Phi(q), |\Phi(p) - d(p,q)|) = 0$ *and,*

(ii) $\sup_{p\in \mathrm{S}_x(T)} |\Phi(p) - \inf_{q\in \mathrm{S}_x(T)} \min(\Phi(q) \dotplus d(p,q))| = 0.$

Also, since $\mathrm{S}_x(T)$ is complete, any continuous function $\Phi : \mathrm{S}_x(T) \to [0,1]$ (with respect to the logic topology) if it satisfies the expressions $(i), (ii)$, or equals 1, then it is a distance function.

Definition 2.25. We call a predicate which is also a distance function a *distance predicate.*

Theorem 2.26. *Let P be a 0-definable x-predicate. Then P is a distance predicate if and only if $\dot{P}$ is a distance function.*

Proof. Suppose $\dot{P}$ is a distance function. If $\dot{P} = 1$, then $P = 1$, which is clearly a distance predicate. Therefore, we assume $\dot{P} \neq 1$. Let $K = \dot{P}^{-1}(0)$ and $A = \iota_x^{-1}(K)$. Then, we have that $\dot{P} = d(\cdot, K)$. Consider $\mathbb{M} \succcurlyeq \boldsymbol{M}$ a monster model (all we need is $\aleph_0$-saturation and $\aleph_0$-strong homogeneity). Let Q be the 0-definable x-predicate on $\mathbb{M}_x$ corresponding to P, and let $B = (\iota_x^{\mathbb{M}})^{-1}(K)$, which is nonempty because K is nonempty and $\iota_x^{\mathbb{M}}$ is a surjection. We have $\iota_x^{\mathbb{M}}(B) = K$ and, by Proposition 2.13, for any element $b \in B$ we have that all the realizations of $\mathrm{tp}(b)$ are in B.

For $a, b \in \mathbb{M}_x$, by the definition of the metric topology on $\mathrm{S}_x(T)$, we have that $d^{\mathbb{M}}(a,b) \geq d(\mathrm{tp}(a), \mathrm{tp}(b))$. As a consequence we obtain $d^{\mathbb{M}}(a,B) \geq d(\mathrm{tp}(a), K)$. On the other hand if $q(x) \in K$, then there exists $c \in B$, with $c \models q(x)$, such that $d(\mathrm{tp}(a), q(x)) = d^{\mathbb{M}}(a,c)$ (this is because first we can find $e, f \in \mathbb{M}_x$ with $e \models \mathrm{tp}(a)$, $f \models q(x)$ such that $d^{\mathbb{M}}(e,f) = d(\mathrm{tp}(a), q(x))$ and then $c = h(f)$ for any automorphism of $\mathbb{M}$ with $h(e) = a$). Hence $d^{\mathbb{M}}(a,B) = d(\mathrm{tp}(a), K) = \dot{P}(\mathrm{tp}(a)) = Q(a)$, and we obtain that Q is a distance function, resulting that it satisfies E_1, E_2. But this implies that P also satisfies E_1, E_2, therefore it's a distance function and it must be with respect to A.

For the second part of the theorem, suppose P is a 0-definable distance predicate. Using the monster model $\mathbb{M}$, previously defined, we define Q to be the 0-definable x-predicate corresponding to P. If $P = 1$, then $Q = 1$. If P satisfies E_1, E_2, then Q also satisfies E_1, E_2. In both cases, Q is a distance predicate. Following the same reasoning as in the previous paragraph we prove $\dot{Q} = \dot{P}$ is a continuous distance function on $\mathrm{S}_x(T)$. □

Corollary 2.27. *Let $A \subseteq M_x$ be a 0-definable set. Then $d(\cdot, A)$ is a 0-definable distance predicate and $(\dot{d(\cdot, A)})$ is a continuous distance function. Conversely if P is a 0-definable distance predicate, then $P^{-1}(0)$ is a 0-definable set.*

Remark 2.28. *If $A \subseteq M_x$ is a 0-definable set then there is an unique 0-definable $K \subseteq \mathrm{S}_x(T)$ such that $\iota_x^{-1}(K) = A$. To see this, suppose K_1, K_2 are distinct 0-definable* *subsets of* $\mathrm{S}_x(T)$ *with this property. Then* $d(\cdot, K_1)$ *and* $d(\cdot, K_2)$ *are two distinct continuous functions* $\mathrm{S}_x(T) \to [0,1]$ *but with the same corresponding 0-definable x-predicate* $d^{\boldsymbol{M}}(\cdot, A)$*, contradiction.*

Using Corollary 2.27, we obtain that the 0-definable sets between elementary structures are in an unique correspondence:

Proposition 2.29. *Let* $\boldsymbol{N}$ *be an* $\mathcal{L}$*-structure, such that* $\boldsymbol{M} \preccurlyeq \boldsymbol{N}$*, and* $A \subseteq M_x$ *be 0-definable. Then there is an unique 0-definable* $B \subseteq N_x$ *such that* $d^{\boldsymbol{N}}(\cdot, B)$ *is the 0-definable x-predicate corresponding to* $d^{\boldsymbol{M}}(\cdot, A)$*. Moreover* $A = B \cap N_x$*, and* $B \neq \emptyset$ *if and only if* $A \neq \emptyset$*.*

Proof. The unicity is obvious. From $(\boldsymbol{M}, d^{\boldsymbol{M}}(\cdot, A)) \preccurlyeq (\boldsymbol{N}, d^{\boldsymbol{N}}(\cdot, B))$ we have that

$$d^{\boldsymbol{M}}(\cdot, A) \neq 1 \Longleftrightarrow d^{\boldsymbol{N}}(\cdot, B) \neq 1,$$

and consequently $A \neq \emptyset \Longleftrightarrow B \neq \emptyset$. □

Remark 2.30. *The 0-definable sets, being zero sets of continuous 0-definable predicates, are closed sets.*

Definition 2.31. Let $S \subseteq M$ be a parameter set. A subset $A \subseteq M_x$ is called an S-*definable* set if A is a 0-definable set over the model $\boldsymbol{M}_S$. We also call A a *definable set* if it is an M-definable set.

We conclude this subsection with a few additional properties of definable sets. However, we first need the following lemma:

Lemma 2.32. *[BBHU08, Proposition 2.10] Let* X *be any set and* $f, g : X \to [0,1]$ *with the property that*

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in X, f(x) \leq \delta \implies g(x) \leq \varepsilon.$$

Then there exists an increasing and continuous function $\alpha : [0,1] \to [0,1]$ *such that* $\alpha(0) = 0$ *and* $\forall x \in X,\ g(x) \leq \alpha(f(x))$*.*

Proposition 2.33. *[BBHU08, Proposition 9.19, Remark 9.20] Let* $A \subseteq M_x$*. The following are equivalent:*

(i) A is a 0-definable set;

(ii) There is an 0-definable x-predicate P such that $A = P^{-1}(0)$ *and*

$$\forall \varepsilon > 0, \exists \delta > 0, \forall a \in M_x,\ P(a) \leq \delta \implies d^{\boldsymbol{M}}(a, A) \leq \varepsilon. \tag{2.1}$$

Moreover if A is the zero-set of a 0-definable predicate P, then A is a 0-definable set if and only if P satisfies statement (2.1)*.*

Proof. The direction $(i) \to (ii)$ is trivial. Suppose (ii). Then, by Lemma 2.32, let $\alpha : [0,1] \to [0,1]$ be an increasing and continuous function satisfying $\alpha(0) = 0$ and $d(a, A) \leq \alpha(P(a))$ for all $a \in M_x$. It remains to observe that $d^{\boldsymbol{M}}(a, A) = \inf_{b \in M_x} \big(\alpha(P(b)) + d^{\boldsymbol{M}}(a, b)\big)$ for all $a \in M_x$.

For the moreover part, if A is a 0-definable set and the zero set of P, then P and $d^{\boldsymbol{M}}(\cdot, A)$ have the same zero set. By Proposition 2.13 item (ii), we obtain that P satisfies statement (2.1). □

Remark 2.34. *[BBHU08, Proposition 9.17] Let $A \subseteq M_x$ be a 0-definable set, y be a multivariable of the same sort as x, distinct from x, and P be a 0-definable xy-predicate. Since P is uniformly continuous applying Lemma 2.32 there exists a function α, with the respective properties, such that $|P(a, b) - P(a, c)| \leq \alpha(d^{\boldsymbol{M}}(b, c))$ for all $a, b, c \in M_x$. Then we have the equalities*

$$\inf_{b \in A}\{P(a, b)\} = \inf_{b \in M_x} (P(a, b) + \alpha(d^{\boldsymbol{M}}(b, A))),$$
$$\sup_{b \in A}\{P(a, b)\} = \sup_{b \in M_x} (P(a, b) - \alpha(d^{\boldsymbol{M}}(b, A))).$$

Thus, the infimum (supremum) of P over A is a 0-definable x-predicate, i.e. for $a \in M_x$ we have that $Q(a) := \inf_{b \in A}\{P(a, b)\}$ ($Q(a) := \sup_{b \in A}\{P(a, b)\}$) is a 0-definable x-predicate. If R is a relation symbol for P, then the notation for the $\mathcal{L}$-formula defining the predicate $\inf_{b \in A}\{P(a, b)\}$ is $\inf_{y \in A} R(x, y)$.

Remark 2.35. *Consider $(\boldsymbol{N}, P, R^{\boldsymbol{N}}) \succcurlyeq (\boldsymbol{M}, Q, R^{\boldsymbol{M}})$ where P, Q are distance 0-definable x-predicates that define the 0-definable sets A and B, respectively, and R is an xy-relation symbol, for some y multivariable. Then, by the previous remark, the 0-definable predicate $(\inf_{y \in A} R(x, y))^{\boldsymbol{M}}$ in $\boldsymbol{M}$ corresponds to the 0-definable predicate $(\inf_{y \in B} R(x, y))^{\boldsymbol{N}}$ in $\boldsymbol{N}$.*

Finally, we present operations and constructions which preserve the definability of sets.

Proposition 2.36. *Let $y = (y_1, \ldots, y_n)$ be a finite multivariable. We have the following:*

(i) If $A, B \subseteq M_x$ are 0-definable, then $A \cup B$ is 0-definable.

(ii) If $A \subseteq M_x$ and $B \subseteq M_y$ are 0-definable sets, then $A \times B$ is 0-definable.

(iii) If $C \subseteq M_{xy}$ is 0-definable then the projection of C in M_x is 0-definable.

(iv) If $A \subseteq M_x$ is 0-definable and $t_1, \ldots, t_n$ are $\mathcal{L}(x)$-terms, then $(t_1, \ldots, t_n)^{\boldsymbol{M}}(A)$ is a 0-definable subset of M_y.

Proof. (i) The set $A \cup B$ is defined by $\min(d(x, A), d(x, B))$.
(ii) The set $A \times B$ is defined by $\max(d(x, A), d(y, B))$.
(iii) The projection of C on M_x is defined by $\inf_y d(xy, C)$.
(iv) The set $(t_1, \ldots, t_n)^{\boldsymbol{M}}(A)$ is defined by $\inf_{x \in A} \max\left(d(y_1, t_1), \ldots, d(y_n, t_n)\right)$. □

Definable functions

Throughout this section x is a finite multivariable and y is a single variable.

In first-order discrete logic a function $f : M_x \to M_y$ is said to be 0-definable in $\boldsymbol{M}$ if there exists an $\mathcal{L}(x, y)$-formula φ such that for all $a \in M_x, b \in M_y$ we have $f(a) = b \Leftrightarrow \varphi^{\boldsymbol{M}}(a, b)$. A natural generalization of this equivalence to continuous logic is $d^{\boldsymbol{M}}(f(a), b) = P(a, b)$ where P is some 0-definable predicate. As with the definable sets, we state a lemma which helps us decide if a definable predicate defines a function.

Lemma 2.37. *Let X, Y be complete metric spaces and $f : X^n \to Y$ be a function. Consider $P : X^n \times Y \to [0, 1]$ such that $P(a, b) = d(f(a), b)$ for all $a \in X^n, b \in Y$. Then P satisfies*

$$\text{(F1)} \quad \sup_{a \in X^n} \inf_{b \in Y} P(a, b) = 0 \textit{ and}$$
$$\text{(F2)} \quad \sup_{a \in X^n} \sup_{b \in Y} \sup_{c \in Y} |d_Y(b, c) - P(a, b)| \dot{-} P(a, c) = 0.$$

Conversely, if $Q : X^n \times Y \to [0, 1]$ satisfies F1 *and* F2*, then the zero set of Q represents the graph of a function f and $Q(a, b) = d_Y(f(a), b)$ for all $a \in X^n$ and $b \in Y$.*

Definition 2.38. We call $f : M_x \to M_y$ a 0-*definable function* if $d^{\boldsymbol{M}}(f(\cdot), \cdot) : M_x \times M_y \to [0, 1]$ is a 0-definable predicate. To specify the domain we also say f *is a* 0-*definable function on* M_x.

By Lemma 2.37, if a 0-definable xy-predicate P satisfies F1, F2, then the zero set of P is the graph of a 0-definable function f and $P(a, b) = d^{\boldsymbol{M}}(f(a), b)$ for every $a \in M_x$ and $b \in M_y$.

Proposition 2.39. *Let $\boldsymbol{M} \preccurlyeq \boldsymbol{N}$. Then the 0-definable functions of $\boldsymbol{M}$ are in a bijective correspondence with the 0-definable functions of $\boldsymbol{N}$: for each 0-definable function $f : M_x \to M_y$ exists an unique 0-definable function $g : N_x \to N_y$ such that $g|_{M_x} = f$.*

Proof. This is easily proved using Lemma 2.37. □

We also call g the 0-definable function corresponding to f. In particular we have $d^{\boldsymbol{N}}(g(\cdot), \cdot)$ is the corresponding 0-definable predicate to $d^{\boldsymbol{M}}(f(\cdot), \cdot)$.

Remark 2.40. *When adding a function symbol to a language we also need to add a modulus of continuity. Note that if f is defined by the 0-definable predicate P, then we can take for f to have the same modulus of continuity as P. Furthermore, the modulus of continuity of P does not depend on $\boldsymbol{M}$, but on the cauchy sequence of formulas defining P.*

Theorem 2.41. *Let $\boldsymbol{M} \preccurlyeq \boldsymbol{N}$ and f, g be two* 0*-definable functions on M_x, N_x respectively. Then $(\boldsymbol{M}, f) \preccurlyeq (\boldsymbol{N}, g)$ if and only if g is the 0-definable function corresponding to f.*

Proof. If $(\boldsymbol{M}, f) \preccurlyeq (\boldsymbol{N}, g)$ then clearly $g|_{M_x} = f$ and g is the 0-definable function corresponding to f. For the converse, suppose g is the 0-definable function corresponding to f, and let α and R be distinct symbols from $\mathcal{L}$ such that $\alpha^{\boldsymbol{M}} = f$ and $R^{\boldsymbol{M}} = d^{\boldsymbol{M}}(f(\cdot), \cdot)$. Clearly $(\boldsymbol{M}, R^{\boldsymbol{M}}) \preccurlyeq (\boldsymbol{N}, R^{\boldsymbol{N}})$. Let $T = \mathrm{Th}(\boldsymbol{M})$ and let (φ_n) define $R^{\boldsymbol{M}}$ such that $\boldsymbol{M} \models \sup |\varphi_n - R| \leq \frac{1}{n+1}$ for every n. We denote

$$T' = T \cup \{\inf_{xy} |\varphi_n - R| \leq \frac{1}{n+1} : n\} \cup \{\sup_{xy} |d(\alpha(x), y) - R(x, y)| = 0\},$$

which is a set of conditions in $\mathcal{L} \cup \{R, \alpha\}$. Clearly $(\boldsymbol{M}, R^{\boldsymbol{M}}, f), (\boldsymbol{N}, R^{\boldsymbol{N}}, g) \models T'$.

Let ψ be an $(\mathcal{L} \cup \{\alpha\})(z)$-formula, for some multivariable z. Then by Proposition 1.48 we have that ψ is approximable by unnested formulas (see Definition 1.47). In particular, for $\varepsilon > 0$ exists an unnested formula ψ_ε such that $\models \sup |\psi - \psi_\varepsilon| \leq \varepsilon$. Using the fact that $T' \models \sup_{xy} |d(\alpha(x), y) - R(x, y)| = 0$, we can replace the atomic instances of the form $d(\alpha(x), y)$ with $R(x, y)$ and obtain an $\mathcal{L} \cup \{R\}$ formula ψ'_ε,

$$T' \models \sup_z |\psi_\varepsilon - \psi'_\varepsilon| = 0.$$

But then for every $c \in M_z$, we have $|\psi^{\boldsymbol{M}}(c) - (\psi'_\varepsilon)^{\boldsymbol{M}}(c)| \leq \varepsilon$, $|\psi^{\boldsymbol{N}}(c) - (\psi'_\varepsilon)^{\boldsymbol{N}}(c)| \leq \varepsilon$. Since $(\boldsymbol{M}, R^{\boldsymbol{M}}) \preccurlyeq (\boldsymbol{N}, R^{\boldsymbol{N}})$, it follows that $(\psi'_\varepsilon)^{\boldsymbol{M}}(c) = (\psi'_\varepsilon)^{\boldsymbol{N}}(c)$. Sending $\varepsilon \to 0$ implies $\psi^{\boldsymbol{M}}(c) = \psi^{\boldsymbol{N}}(c)$, from which we conclude that $(\boldsymbol{M}, f) \preccurlyeq (\boldsymbol{N}, g)$. □

We also have the following interesting result:

Proposition 2.42. *Let f be a 0-definable function on M_x defined by the 0-definable predicate $P = R^{\boldsymbol{M}}$. Then for any $p(x) \in \mathrm{S}_x(\mathrm{Th}(\boldsymbol{M}, P))$ there is an unique xy-type extending $p(x) \cup \{R(x, y) = 0\}$.*

Proof. Let $q(x, y)$ and $q'(x, y)$ be two types extending $p(x) \cup \{R(x, y) = 0\}$. Let $\boldsymbol{N} \succcurlyeq \boldsymbol{M}$ (a monster model extension) with $(a, b), (a', b') \in N_x$ such that $(a, b) \models q$ and $(a', b') \models q'$. Then there exists and automorphism h of $\boldsymbol{N}$ such that $h(a) = a'$. And then since $R^{\boldsymbol{N}}(a, b) = 0, R^{\boldsymbol{N}}(a', b') = 0$, with b and b' being the unique elements for which the equalities hold, we obtain $h(b) = b'$. Hence, obviously, $q = q'$. □

We continue with a discussion about how we interpret the 0-definable functions on the type spaces.

Notation: Let w, z be multivariables of the same sort. Then, we define the function $t^w_z : \mathrm{S}_w(T) \to \mathrm{S}_z(T)$ such that $t^w_z(p(w))$ is the z-type obtained by replacing w with z in p. Let $\Phi : \mathrm{S}_w(T) \to [0, 1]$ be a continuous function. To emphasize the domain $\mathrm{S}_w(T)$, we write $\Phi(w)$. Then, we write $\Phi(z)$ for the function $\Phi(w) \circ t^w_z$, so that $\Phi(z) : \mathrm{S}_z(T) \to [0, 1]$. Moreover, given any $u \supseteq w$, by abuse of notation, we will also denote by $\Phi(w)$ the function $\Phi(w)(\cdot|_w) : \mathrm{S}_u(T) \to [0, 1]$.

Let f be a 0-definable x-function defined by the 0-definable xy-predicate $R^{\boldsymbol{M}}$ and let $\Phi = \dot{R}$. Then $\boldsymbol{M} \models \sup_x \inf_y R(x, y) = 0$ and $\boldsymbol{M} \models \sup_x \sup_y \sup_z |d(y, z) - R(x, y)| - R(x, z) = 0$. Then, according to the ideas mentioned in subsection "The logic mechanism on types" of Section 1.4 "Type Spaces", and subsection "More on definable predicates" of section 2.1 "Definability", we obtain that:

$$\text{(F1)}\ \sup_x \sup_y \Phi = 0 \text{ and}$$

$$\text{(F2)}\ \sup_x \sup_x \sup_z |d(\dot{y}, z) - \Phi(xy)| - \Phi(xz) = 0.$$

The functions $d(\dot{y}, z), \Phi(xy), \Phi(xz)$ have domain $\mathrm{S}_{xyz}(T)$ and $d(\dot{y}, z)(p) = d(\dot{y}, z)(p|_{y,z})$. From $(F1)$ and Proposition 2.17 we obtain that for any $p(x) \in \mathrm{S}_x(T)$ exists an $q \in \mathrm{S}_{xy}(T)$ with $p \subseteq q$ and $\Phi(q) = 0$. Suppose there exists another $q'(xy) \supseteq p(x)$ such that $\Phi(q') = 0$. Let and xyz-type $r(x, y, z) \supseteq q(x, y) \cup t^{xy}_{xz}(q')$ (such an r clearly exists). Then by F2 we obtain $(d(y, z))^r = 0$, but this obviously implies $q = q'$. Hence every $p(x) \in \mathrm{S}_x(T)$ has an unique extension to a $q(x, y) \in \mathrm{S}_{xy}(T)$ such that $\Phi(q) = 0$. By Theorem 2.19, $p(x)$ has an unique extension to a type in $\mathrm{S}_x(\mathrm{Th}(\boldsymbol{M}, R^{\boldsymbol{M}}))$, by abuse of notation we denote this extension by $p(x)$ as well (similarly for $q(x, y)$). Observe that $R(x, y) = 0 \in q(x, y)$. This implies that $p(x) \cup \{R(x, y) = 0\}$ has an unique extension to an xy-type. Thereby reproving Proposition 2.42.

Remark 2.43. *We observe that working with formulas defined on type spaces allows us to avoid the use of monster model arguments (as in the proof of Proposition 2.42 in the previous paragraph), which, for instance, were employed in the context of definable sets and in Proposition 2.41.*

Finally, we construct the function $\dot{f} : \mathrm{S}_x(T) \to \mathrm{S}_y(T)$ where $\dot{f}(p(x)) = q|_y$ with q the unique xy-type extending p and $\Phi(q) = 0$. As expected, we have:

$$\begin{array}{ccc} M_x & \xrightarrow{f} & M_y \\ {\scriptstyle \iota_x}\downarrow & & \downarrow{\scriptstyle \iota_y} \\ \mathrm{S}_x(T) & \xrightarrow{\dot{f}} & \mathrm{S}_y(T). \end{array}$$

Remark 2.44. *We could call $\dot{f}$ a 0-definable function on the type space $\mathrm{S}_x(T)$. However, this would require to also find an intrinsic characterization of 0-definable functions on $\mathrm{S}_x(T)$ solely in terms of type spaces and formulas defined on them. At present, this seems a bit complicated to do, as notation quickly becomes cumbersome and there is still a lack of results and understanding of the interplay between all the type spaces and the formulas defined on them. The author leaves this matter for future research.*

We end this section with some basic results on 0-definable functions. Let z, v be finite multivariables.

Remark 2.45. *The definition of a 0-definable function $M_x \to M_y$ (with y a single variable) extends naturally to functions $M_x \to M_z$. We call a function $f : M_x \to M_z$ 0-definable if the predicate $d^{\boldsymbol{M}}(f(\cdot), \cdot) : M_x \times M_z \to [0, 1]$ is 0-definable. The corresponding Proposition 2.39 and Theorem 3.9 continue to hold in this more general setting.*

Proposition 2.46. *We have the following:*

(i) For any $\mathcal{L}(x)$-term t we have that $t^{\boldsymbol{M}}$ is 0-definable.

(ii) If $f : M_x \to M_z, g : M_x \to M_z$ are 0-definable functions then $(f, g) : M_x \to M_z$ is a 0-definable function.

(iii) If $f : M_x \to M_z, g : M_z \to M_v$ be 0-definable functions, then gf is a 0-definable function.

(iv) If $A \subseteq M_x$ is a 0-definable set and $f : M_x \to M_z$ is a 0-definable function, then $f(A)$ is a 0-definable set.

Proof. (i) The map $t^{\boldsymbol{M}}$ is defined by $d(t(x), y)$.
(ii) The map $(f, g) : M_x \to M_z \times M_z$ is defined by $\max(d(f(x), z), d(g(x), z'))$, for z' a multivariable of the same type as z.
The proof of (iii) follows a similar strategy as proof of Proposition 2.41. We use the fact that $d(gf(x), v)$ is approximable by unnested formulas in $\mathcal{L} \cup \{f, g\}$, and then replace the atomic formulas with the 0-definable predicates defining f and g, such that in the end to obtain an $\mathcal{L}$-formula approximating $d(gf(x), v)$.
(iv) Obviously $f(A)$ is defined by $\inf_{x \in A} d(f(x), z)$, which is a 0-definable predicate over $\mathcal{L}$. □

Remark 2.47. *Observe that by item (iii) of Proposition 2.46, even if we add some 0-definable functions to the language then in the expanded model of $\boldsymbol{M}$ the set of 0-functions is the same as the set of 0-definable functions in $\boldsymbol{M}$ as an $\mathcal{L}$-structure.*

Extending the language

Until now, we have worked with a complete $\mathcal{L}$-theory T. However, definability concepts can similarly be introduced in the context of incomplete theories. Hence, let T be any theory. For example, using a cauchy sequence of formulas $(\varphi_n(x))$ (the same as Remark 2.5 for any theory) we can define a 0-definable x-predicate on any model $\boldsymbol{M} \models T$ by being the limit of $(\varphi^{\boldsymbol{M}})$. Then we can add a relation symbol R to the language $\mathcal{L}$ and add to the theory T the axioms $\{\sup_x |R(x) - \varphi_{n_\varepsilon}(x)| \leq \varepsilon : \varepsilon > 0\}$, with n_ε chosen accordingly. By slight abuse of terminology, we may refer to R itself as a 0-definable predicate.

Moreover, a 0-definable function can clearly be defined by any 0-definable predicate xy-symbol R which satisfies conditions F1 and F2 i.e.

$$\begin{aligned} T \cup \{\sup_x |R(xy) - \varphi_{n_\varepsilon}(xy)| \leq \varepsilon : \varepsilon > 0\} &\models \sup_x \inf_y R(x, y) = 0, \\ &\models \sup_x \sup_y \sup_z |d(y, z) - R(x, y)| - R(x, z) = 0. \end{aligned}$$

Then, we add to the language $L \cup \{R\}$ the symbol f and to the theory $T \cup \{\sup_x |R(x) - \varphi_{n_\varepsilon}(x)| \leq \varepsilon : \varepsilon > 0\}$ the axiom $\sup_x \sup_y |d(f(x), y) - R(x, y)| = 0$. Similarly we may refer to the symbol f as a 0-definable function. Note that it is not necessary to have the symbol R in the language to define f. It's enough to add to T the axioms $\{\sup_x \sup_y |d(f(x), y) - \varphi_{n_\varepsilon}(x, y)| \leq \varepsilon : \varepsilon > 0\}$.

Definition 2.48. Let $\mathcal{L}_0 \subseteq \mathcal{L}$ be a language and T_0 be a $\mathcal{L}_0$-theory. We call T a *conservative extension* of T_0 if $T \models E$ if and only if $T_0 \models E$ for all closed $\mathcal{L}$-conditions E. An $\mathcal{L}(x)$-formula $\varphi(x)$ is called *definable in* T *over* $\mathcal{L}_0$ if for every $\varepsilon > 0$ there exists an $\mathcal{L}_0$-formula $\psi_\varepsilon(x)$ such that $T \models \sup_x |\varphi(x) - \psi_\varepsilon(x)| \leq \varepsilon$. If T is a conservative extension of T_0, we also call T an *extension by definitions* of T_0 if every $\mathcal{L}$-formula is defined in T over $\mathcal{L}_0$.

We say that a relation symbol $R \in \mathcal{L}$ ($f \in \mathcal{L}$) is *definable in* T *over* $\mathcal{L}_0$ if $R(x)$ ($d(f(x), y)$) is definable in T over $\mathcal{L}_0$.

The following is Proposition 9.28 from [BBHU08], where they use Theorem 1.41 to provide a complex, but short and elegant proof.

Theorem 2.49. *Let $\mathcal{L}_0 \subseteq \mathcal{L}$ be a language and T_0 be an $\mathcal{L}_0$-theory. If T is a conservative extension of T_0 and every non-logical symbol from $\mathcal{L}$ is defined in T over $\mathcal{L}_0$, then T is an extension by definitions of T_0.*

Proof. Let φ be an $\mathcal{L}(x)$-formula and fix $\varepsilon > 0$. By Proposition 1.48 there exists an unnested $\mathcal{L}$-formula φ_ε such that $\models \sup_x |\varphi - \varphi_\varepsilon| \leq \varepsilon$. Using the fact that formulas are uniformly continuous with respect to subformulas, we can replace the atomic formulas of φ_ε which are of the form $R(z)$ and $d(f(z'), z'')$, for some multivariables $z, z'z,''$, with $\mathcal{L}_0$-formulas which approximate $R(z)$ and $d(f(z'), z'')$, in order to obtain over T an $\mathcal{L}_0$-formula approximating φ_ε. This is possible because each non-logical symbol from $\mathcal{L}$ is defined in T over $\mathcal{L}_0$. It follows that there exists an $\mathcal{L}_0$-formula φ'_ε such that $T \models \sup_x |\varphi_\varepsilon - \varphi'_\varepsilon| \leq \varepsilon$, and respectively $T \models \sup_x |\varphi - \varphi'_\varepsilon| \leq 2\varepsilon$. Concluding that T is an extension by definitions of T_0. $\square$

Finally, just as in first-order classical logic, if T is an extension by definitions of T_0 then any model $\boldsymbol{M}_0 \models T_0$ can be uniquely extended to a model $\boldsymbol{M} \models T$ [BBHU08, Corollary 9.31].

Remark 2.50. *Suppose $\boldsymbol{M} \preccurlyeq \boldsymbol{N}$. Let $\{P_i\}_{i\in I}, \{f_j\}_{j\in J}$ be a collection of 0-definable predicates and 0-definable functions over $\boldsymbol{M}$, and let $\{Q_i\}_{i\in I}, \{g_j\}_{j\in J}$ be a collection of 0-definable predicates and 0-definable functions over $\boldsymbol{N}$. Then using Theorem 2.49 it is immediate to prove that $(\boldsymbol{M}, \{P_i\}_{i\in I}, \{f_j\}_{j\in J}) \preccurlyeq (\boldsymbol{N}, \{Q_i\}_{i\in I}, \{g_j\}_{j\in J})$ if and only if P_i is the corresponding 0-definable predicate to Q_i, and g_j is the corresponding 0-definable function to f_j, for each $i \in I$ and $j \in J$.*

2.2 Algebraicity

Throughout this subsection let T be a complete $\mathcal{L}$-theory, $\boldsymbol{M}$ be a model of T, $A \subseteq M$ be a parameter set, and $x = (x_1, \ldots, x_n)$ be a finite multivariable.

In classical model theory, a finite definable set in a model remains definable in any elementary extension. The elements of such sets are called algebraic, and if the definable set is a singleton, its element is called definable. A similar situation occurs in continuous model theory where "finite" is replaced with "compact".

Theorem 2.51. *Let $C \subseteq M_x$ be an A-definable set in $\boldsymbol{M}$. Then C is an A-definable set in any elementary extension of $\boldsymbol{M}$ if and only if C is compact.*

Proof. Suppose C remains A-definable in any elementary extension of $\boldsymbol{M}$ and let R be a relation symbol with $R^{\boldsymbol{M}} = d^{\boldsymbol{M}}(\cdot, C)$. Take $(\boldsymbol{N}, R^{\boldsymbol{N}}) \succcurlyeq (\boldsymbol{M}, R^{\boldsymbol{M}})$ a $|C|^+$-saturated model. Clearly $C = (R^{\boldsymbol{N}})^{-1}(0)$. Then

$$p(x) = \{R(x) = 0\} \cup \{d(x, c) \geq \varepsilon : c \in C\}$$

is not realizable in $\boldsymbol{N}$. Consequently $p(x)$ is not finitely satisfiable in $\boldsymbol{N}$, which implies that there exists $c_1, \ldots, c_n \in C$ such that

$$\boldsymbol{N} \models \sup_{x \in C} \min(d(x, c_1), \ldots, d(x, c_n)\} \leq \varepsilon,$$

notation according to Remark 2.34. Therefore, for every $\varepsilon > 0$ we have that C has a finite ε-net. Since C is closed, complete and totally bounded, results that it is compact [Roy10, Theorem 16, p. 199].

Conversely, suppose C is compact. Let P be the A-definable predicate in $\boldsymbol{M}$ defining C, and let Q be the corresponding A-definable predicate in $\boldsymbol{N} \succcurlyeq \boldsymbol{M}$. Then we have that for every $\varepsilon > 0$ exists m and $c_1, \ldots, c_m \in C$ such that

$$\boldsymbol{M} \models \sup_{x \in C} \min(d(x, c_1), \ldots, d(x, c_m)) \leq \varepsilon.$$

Hence, $\boldsymbol{N}$ also satisfies the same condition with C replaced by $Q^{-1}(0)$. Since $C \subseteq Q^{-1}(0)$, we obtain that C is dense in $Q^{-1}(0)$, and trivially $C = Q^{-1}(0)$. $\square$

Thus, compact definable sets are preserved under any elementary extension. Furthermore, by Proposition 2.51, it is straightforward to prove that compact definable sets are also preserved under elementary substructures.

Definition 2.52. We call $a \in M_x$ *definable over* A (or A-*definable*) if $\{a\}$ is an A-definable set. We call $a \in M_x$ *algebraic over* A (or A-*algebraic*) if there exists a compact A-definable set containing a. If $A = \emptyset$ then we use the notation 0-definable and 0-algebraic in place of $\emptyset$-definable and $\emptyset$-algebraic. Similarly when $A = M$, we simply say *definable* and *algebraic*.

Proposition 2.53. *[BBHU08, Proposition 10.2] An element $a = (a_1, \ldots, a_n) \in M_x$ is A-definable/algebraic if an only if a_m is A-definable/algebraic for every m.*

Proof. We prove the algebraic case. If C is a compact A-definable set containing a then a_1 (and similarly for $a_2, \ldots, a_n$) is included in C_1 (the x_1 projection of C) which is defined by $\inf_{x_2,\ldots,x_n} d(x, C)$. Conversely, suppose $a_1, \ldots, a_n$ are A-algebraic and that for each m, there exists a compact A-definable set C_m such that $a_m \in C_m$. Then a is contained in $\prod_m C_m$, which is defined by $\max(d(x_1, C_1), \ldots, d(x_n, C_n))$. $\square$

Definition 2.54. By $\mathrm{dcl}_{\boldsymbol{M}}(A) \subseteq M$ we denote the collection of all A-definable elements, and by $\mathrm{acl}_{\boldsymbol{M}}(A) \subseteq M$ we denote the collection of all A-algebraic elements. By Theorem 2.51, for any $\boldsymbol{N} \succcurlyeq \boldsymbol{M}$ we have $\mathrm{dcl}_{\boldsymbol{M}}(A) = \mathrm{dcl}_{\boldsymbol{N}}(A)$ and $\mathrm{acl}_{\boldsymbol{M}}(A) = \mathrm{acl}_{\boldsymbol{N}}(A)$. Hence we can simply write $\mathrm{dcl}(A)$ and $\mathrm{acl}(A)$. Moreover, it is clear that $\mathrm{dcl}(A) \subseteq \mathrm{acl}(A)$.

Proposition 2.55. *We have the following:*

(i) The interpretation of each constant symbol of $\boldsymbol{M}$ is 0*-definable.*

(ii) For every $\mathcal{L}(x)$-term t if $c \in M_x$ is 0-definable, then so is $t^{\boldsymbol{M}}(c)$.

(iii) If $\boldsymbol{M}'$ is an expansion by definitions of $\boldsymbol{M}$, then $\mathrm{dcl}^{\boldsymbol{M}}(\emptyset) = \mathrm{dcl}^{\boldsymbol{M}'}(\emptyset)$.

Proof. (i) For a constant symbol $c \in \mathcal{L}$, we have that $c^{\boldsymbol{M}}$ is 0-definable by $d(c, y)$.
(ii) We have that $t^{\boldsymbol{M}}(c)$ is 0-definable by $\inf_{x \in \{c\}} d(t(x), y)$
(iii) This is a consequence of Theorem 2.49. □

Proposition 2.56. *We have the following:*

1. *For every $\mathcal{L}(x)$-term t if $a \subseteq M_x$ 0-algebraic, then so is $t^{\boldsymbol{M}}(a)$.*
2. *If $\boldsymbol{M}'$ is an expansion by definitions of $\boldsymbol{M}$, then* $\mathrm{acl}^{\boldsymbol{M}}(\emptyset) = \mathrm{acl}^{\boldsymbol{M}'}(\emptyset)$.

Proof. (i) Let $a \in C \subseteq M_x$, with C a compact 0-definable set. Then, by Proposition 2.36(iv), we have that $t^{\boldsymbol{M}}(C)$ is a compact 0-definable set. Concluding $t^{\boldsymbol{M}}(a)$ is 0-definable.
(ii) This is a consequence of Theorem 2.49. □

Before proving Theorem 2.61, we give a characterization for when a zero set of a definable predicate is a definable set.

Proposition 2.57. *Let P be a 0-definable x-predicate in $\boldsymbol{M}$. Suppose that for every $\varepsilon > 0$ exists $\delta > 0$ such that $P^{-1}([0, \delta])$ has a finite ε-net. Then $P^{-1}(0)$ is a compact definable set.*

Proof. We will prove the condition in Proposition 2.33(ii). Let $A = P^{-1}(0)$. Suppose there exists $\varepsilon > 0$ such that for every $\delta > 0$ exists $a \in M_x$ such that $P(a) \leq \delta$ and $d(a, A) > \varepsilon$. Let (δ_n) be a decreasing sequence to 0 such that $P^{-1}([0, \delta_n])$ has a finite $1/n$-net, and for each n let $a_n \in P^{-1}([0, \delta_n])$ such that $d(a_n, A) > \varepsilon$. Then, by a diagonal argument there exists a cauchy subsequence $(a_{n_m})_m$. Since M_x is complete let $a = \lim_m a_{n_m}$. By continuity of P we have $P(a) = 0$, which means $a \in A$, and by continuity of $d(\cdot, A)$ we obtain $d(a, A) \geq \varepsilon$, contradiction. Therefore, $P^{-1}(0)$ is a definable set. Since it obviously is totally bounded, it follows that it is compact. □

Corollary 2.58. *Let P be a 0-definable x-predicate in $\boldsymbol{M}$. Suppose that for every $\varepsilon > 0$ exists $\delta > 0$ such that $P^{-1}([0, \delta])$ has diameter $\leq \varepsilon$. Then $P^{-1}(0)$ is a 0-definable element.*

We will need the following useful corollary:

Corollary 2.59. *[BBHU08, Exercise 10.8(3)] The element $a \in M_x$ is 0-algebraic if and only if for every $\varepsilon > 0$ exists $\delta > 0$ and a 0-definable predicate $P_\varepsilon(x)$ such that $P_\varepsilon(a) = 0$ and $P_\varepsilon^{-1}([0,\delta])$ has a finite ε-net.*

Proof. The left to right direction is trivial. For the right to left, observe that $P : M_x \to [0,1]$ defined by

$$P(b) = \sum_{n=1}^{\infty} \frac{P_{1/n}(b)}{2^n},$$

for every $b \in M_x$ is a 0-definable predicate satisfying the condition of Proposition 2.57. $\square$

Corollary 2.60. *[BBHU08, Exercise 10.7(3)] The element $a \in M_x$ is 0-definable if and only if for every $\varepsilon > 0$ exists $\delta > 0$ and a 0-definable predicate $P_\varepsilon(x)$ such that $P_\varepsilon(a) = 0$ and $P_\varepsilon^{-1}([0,\delta])$ has diameter $\leq \varepsilon$.*

In classical model theory is very easy to prove $\mathrm{acl}(\emptyset) = \mathrm{acl}(\mathrm{acl}(\emptyset))$. However, in continuous logic, establishing this equality is significantly more challenging. This fact is proved in [WH07, Proposition 2.3] using the properties of a monster model. Here we provide a proof just using formulas and studying the model $\boldsymbol{M}$.

Theorem 2.61. *We have the following:*

(i) $\mathrm{dcl}(\emptyset) = \mathrm{dcl}(\mathrm{dcl}(\emptyset))$;

(ii) $\mathrm{acl}(\emptyset) = \mathrm{acl}(\mathrm{acl}(\emptyset))$.

Proof. We show a proof for item *(ii)* (item *(i)* is left to the reader). Suppose x is a single variable. Let $a \in \mathrm{acl}(\mathrm{acl}(\emptyset))$, and $a \in C$ with $C \subseteq M$ a compact $\mathrm{acl}(\emptyset)$-definable set. Let $P = d(\cdot, C)$ (which is an $\mathrm{acl}(\emptyset)$-definable predicate) and fix $\varepsilon > 0$. Let φ be an $\mathcal{L}(y,x)$-formula and $K \subseteq M_y$ be a compact definable set and $k \in K$ such that

$$\boldsymbol{M} \models \sup_x |P(x) - \varphi(k,x)| \leq \varepsilon, \varphi(k,a) = 0.$$

Then $(\varphi(k,x)^{\boldsymbol{M}})^{-1}([0,2\varepsilon])$ has a finite 4ε-net consisting of m elements, for some m. Consider the following formula

$$\psi(y) = \inf_{x_1,\ldots,x_m} \sup_x \left(\left(2\varepsilon \dot{-} \varphi(y,x)\right) \wedge \left(\min(d(x,x_1),\ldots,d(x,x_m)) \dot{-} 4\varepsilon \right) \right).$$

Then for every $b \in M$, $\boldsymbol{M} \models \psi(b) = 0$ says that $(\phi(b,x)^{\boldsymbol{M}})([0,2\varepsilon])$ has a finite ε'-net for any $\varepsilon' > 4\varepsilon$, in particular has a finite 5ε-net. Observe that $\psi(a)$ holds.

Claim. Let $\alpha(y,x)$ be any $\mathcal{L}$-formula. Then there exists $k_1,\ldots,k_n \in K$ such that for any $k \in K$ there exists an index $m \leq n$ such that $|\alpha(k,x)^{\boldsymbol{M}} - \alpha(k_m,x)^{\boldsymbol{M}}| \leq \varepsilon$ (here $|\cdot|$ is the supremum norm on $\mathrm{C}(M_x,[0,1])$). In particular $(\alpha(k,x)^{\boldsymbol{M}})^{-1}([0,\varepsilon]) \subseteq (\alpha(k_m,x)^{\boldsymbol{M}})^{-1}([0,2\varepsilon])$.

Proof of Claim. The claim is proven easily using the fact that $\alpha^{\boldsymbol{M}}$ is an uniform function and K has a finite ε'-net for any $\varepsilon' > 0$. □

We now examine the formula

$$\gamma(x) = \inf_{y \in K}((\varepsilon + \psi(y)) \vee \varphi(y, x)),$$

and consider the set $T = (\gamma^{\boldsymbol{M}})^{-1}([0, \varepsilon])$. Since K is compact (and clearly complete), for any $b \in T$ exists $k_b \in K$ such that $\psi(k_b) = 0$ and $\varphi(k_b, b) \leq \varepsilon$. Let $K' = K \cap (\psi^{\boldsymbol{M}})^{-1}(0)$. Applying the Claim to φ and K' we obtain $k_1, \ldots, k_n$ such that for any $k \in K'$ exists some $m \leq n$ satisfying

$$(\varphi(k, x)^{\boldsymbol{M}})^{-1}([0, \varepsilon]) \subseteq (\varphi(k_m, x)^{\boldsymbol{M}})^{-1}([0, 2\varepsilon]).$$

From this, it follows that

$$(\gamma(x))^{-1}([0, \varepsilon]) \subseteq \bigcup_{1 \leq m \leq n} (\varphi(k_m, x)^{\boldsymbol{M}})^{-1}([0, 2\varepsilon]).$$

Finally, since for each m, the set $(\varphi(k_m, x)^{\boldsymbol{M}})^{-1}([0, 2\varepsilon])$ has a finite 5ε-net, it follows that $(\gamma(x))^{-1}([0, \varepsilon])$ also has a finite 5ε-net. In summary, we have shown that for every $\varepsilon > 0$, there exists an $\mathcal{L}$-formula $\gamma(x)$ such that $((\gamma)^{\boldsymbol{M}})^{-1}([0, \varepsilon])$ has a finite 5ε-net. By Corollary 2.59, we conclude that a is 0-algebraic. □

As a consequence of the results of this section we can easily obtain the following fundamental characterization of definable and algebraic elements.

Theorem 2.62. *[BBHU08, Exercises 10.7,10.8] Let $a \in M_x$, and $\mathbb{M} \succcurlyeq \boldsymbol{M}$ a monster model extension. We have the following:*

(i) The element a is A-definable if and only if the x-type $\mathrm{tp}(a|A)$ has an unique realization in any elementary extension of $\boldsymbol{M}$.

(ii) The element a is A-definable if and only if $\{a\} = \{h(a) : h \in \mathrm{Aut}(\mathbb{M}|A)\}$.

(iii) The element a is A-algebraic if and only if the x-type $\mathrm{tp}(a|A)$ has the same set of realizations in any elementary extension of $\boldsymbol{M}$.

(iv) The element a is A-algebraic if and only if $\{h(a) : h \in \mathrm{Aut}(\mathbb{M}|A)\}$ is small.

2.3 Topics in Model Theory

An essential topological lemma

We begin with a topological fact that is particularly useful in continuous model theory. Let X be a compact and Hausdorff space, Y be a Hausdorff space, and $\mathcal{F}$ be a family of continuous functions $X \to Y$. Let $X/\sim$ be the quotient set of X where $a, b \in X$ are in equivalence if and only if for all $f \in \mathcal{F}$ we have $f(a) = f(b)$. We also denote $X/\mathcal{F} = X/\sim$ and call it the quotient of X by $\mathcal{F}$. Then every function $f \in \mathcal{F}$ factors uniquely through $X/\mathcal{F}$, that is given the natural projection $\pi : X \to X/\mathcal{F}$ there exists an unique $\tilde{f} : X/\mathcal{F} \to Y$ such that $f = \tilde{f} \circ \pi$. We denote by $\widetilde{\mathcal{F}} = \{\tilde{f} : f \in \mathcal{F}\}$.

Lemma 2.63. *[BYU10, Fact 4.7] The quotient topology on $X/\mathcal{F}$ is the same as the initial topology on $X/\mathcal{F}$ induced by $\widetilde{\mathcal{F}}$. This topology is compact and Hausdorff.*

Proof. This fact has already been proved for $\mathcal{F}$ a singleton in the paragraphs below Proposition 2.13. For any $\mathcal{F}$ the proof is similar. It suffices to observe that the initial topology on $X/\mathcal{F}$ is Hausdorff and contained in the quotient topology on $X/\mathcal{F}$ which is compact, hence they must coincide. $\square$

The back-and-forth method

Throughout this subsection x is a finite multivariable.

Definition 2.64. Let $\boldsymbol{M}, \boldsymbol{N}$ be two $\mathcal{L}$-structures. A partial isomorphism from $\boldsymbol{M}$ to $\boldsymbol{N}$ is a bijective map $\gamma : A \to B$ with $A \subseteq M$, $B \subseteq N$, such that

(i) for each n-ary relation symbol $R \in \mathcal{L}^r$ and $a \in A^n$ we have $R^{\boldsymbol{M}}(a) = R^{\boldsymbol{N}}(\gamma(a))$;

(ii) for each n-ary $f \in \mathcal{L}^f$ and $a \in A^n, b \in A$ we have $f^{\boldsymbol{M}}(a) = b \leftrightarrow f^{\boldsymbol{M}}(\gamma(a)) = \gamma(b)$.

Then, just as in classical logic we define:

A *back-and-forth system* from $\boldsymbol{M}$ to $\boldsymbol{N}$ is a nonempty collection Γ of partial isomorphisms from $\boldsymbol{M}$ to $\boldsymbol{N}$, such that for every $\gamma \in \Gamma$ we have that

(i) for every $a \in M$, there exists $\gamma' \in \Gamma$ extending γ with $a \in \mathrm{dom}(\gamma')$; and

(ii) for every $b \in N$, there exists $\gamma' \in \Gamma$ extending γ with $b \in \mathrm{im}(\gamma')$.

Theorem 2.65. *Let Γ be a back-and-forth system from $\boldsymbol{M}$ to $\boldsymbol{N}$. Then each $\gamma \in \Gamma$ is an elementary map* $\mathrm{dom}(\gamma) \to N$.

Proof. Since formulas are approximable by unnested formulas, it's enough to prove that for every $\gamma \in \Gamma$, unnested $\mathcal{L}(x)$-formula φ, and every $a \in M_x$ we have the equality

$$\varphi^{\boldsymbol{M}}(a) = \varphi^{\boldsymbol{N}}(\gamma(a)). \tag{2.2}$$

This is done by induction on the number of connectives in φ. First, if φ is an atomic unnested formula then (2.2) holds trivially. Moreover the equality is preserved under non-quantifier connectives. Suppose $\varphi = \inf_y \psi$, with y a single variable. Let $r \in [0, 1]$ be such that $\varphi^{\boldsymbol{M}}(a) < r$, for some $a \in \mathrm{dom}(\gamma)_x$. Then there exists $b \in M$ such that $\psi^{\boldsymbol{M}}(a, b) < r$. Let $\gamma' \in \Gamma$ be an extension of γ with $b \in \mathrm{dom}(\gamma')$. By the inductive hypothesis $\psi^{\boldsymbol{N}}(\gamma(a), \gamma(b)) < r$, consequently $\varphi^{\boldsymbol{N}}(\gamma(a)) < r$. It follows that $\varphi^{\boldsymbol{N}}(\gamma(a)) \leq \varphi^{\boldsymbol{M}}(a)$. Similarly is proven $\varphi^{\boldsymbol{N}}(\gamma(a)) \geq \varphi^{\boldsymbol{M}}(a)$. Therefore φ satisfies (2.2). $\square$

Corollary 2.66. *If there is a back-and-forth system between $\boldsymbol{M}$ and $\boldsymbol{N}$, then $\boldsymbol{M} \equiv \boldsymbol{N}$.*

Quantifier Elimination

For this subsection we assume $\mathcal{L}$ has constant symbols. Let T be an $\mathcal{L}$-theory, and x a finite multivariable.

Definition 2.67. An $\mathcal{L}(x)$-formula φ is said to be *approximable in T by quantifier-free formulas* if for every $\varepsilon > 0$ there is a quantifier-free $\mathcal{L}(x)$-formula φ^{qf} such that

$$T \models \sup_x |\varphi - \varphi^{\mathrm{qf}}| \leq \varepsilon.$$

Therefore φ would be equivalent to a definable predicate defined by a limit of quantifier-free formulas, such definable predicates are also called *quantifier-free definable predicates.*

We say T has *quantifier elimination* (QE) if every $\mathcal{L}$-formula is approximable in T by quantifier-free formulas.

Definition 2.68. For a parameter set $A \subseteq \boldsymbol{M}$, we denote by $\langle A \rangle_{\boldsymbol{M}}$ the unique substructure generated by A.

Theorem 2.69. *[BBHU08, Proposition 13.2] Let φ be an $\mathcal{L}(x)$-formula. Then φ is approximable in T by quantifier-free formulas if and only if for all $\boldsymbol{M}, \boldsymbol{N} \models T$ every common substructure A of $\boldsymbol{M}$ and $\boldsymbol{N}$ and all $a \in A_x$ we have $\varphi(a)^{\boldsymbol{M}} = \varphi(a)^{\boldsymbol{N}}$.*

Proof. The left to right direction is trivial. For the right to left direction, consider the function $\dot{\varphi} : \mathrm{S}_x(T) \to [0,1]$. Let $\mathcal{F} = \{\dot{\psi} : \psi \text{ is a quantifier-free } \mathcal{L}(x)\text{-formula}\}$. Denote $\mathrm{S}_x^{\mathrm{qf}}(T) := \mathrm{S}_x(T)/\mathcal{F}$. By Lemma 2.63 we obtain that $\mathrm{S}_x^{\mathrm{qf}}(T)$ is a compact Hausdorff space. Moreover, by Proposition 1.4, the set $\widetilde{\mathcal{F}}$ is dense in $\mathrm{C}(\mathrm{S}_x^{\mathrm{qf}}(T), [0,1])$. Additionally, the natural projection $\pi : \mathrm{S}_x(T) \to \mathrm{S}_x^{qf}(T)$ is a continuous surjection. Therefore, using Lemma 2.14, any continuous function $\mathrm{S}_x(T) \to [0,1]$ which factors through $\mathrm{S}_x^{\mathrm{qf}}(T)$ is equivalent to a quantifier-free definable predicate.

Let $p, q \in \mathrm{S}_x(T)$ such that $\pi(p) = \pi(q)$ and let $\boldsymbol{M}, \boldsymbol{N} \models T$ for which there exists $a \in M_x, a \models p$ and $b \in N_x, b \models q$. Then since a and b satisfy the same quantifier-free formulas it is easy to see that there exists an isomorphism $\gamma : \langle a \rangle_{\boldsymbol{M}} \to \langle b \rangle_{\boldsymbol{N}}$ sending $a \to b$. Finally, using γ and our assumption we obtain that $\varphi(a)^{\boldsymbol{M}} = \varphi(b)^{\boldsymbol{N}}$, which in turn implies $\dot{\varphi}(p) = \dot{\varphi}(q)$. Consequently $\dot{\varphi}$ factors through $\mathrm{S}_x^{\mathrm{qf}}(T)$, proving that it is equivalent to a quantifier-free definable predicate. $\square$

3 Fundamental Theorem of Stability

This chapter begins with an introduction to stability theory in continuous logic. Its central focus is the Fundamental Theorem of Stability (FTS), which establishes the equivalence between the stability of a formula φ and the definability of every φ-type. In the first section, we provide an original and elementary proof of the FTS, which only uses the definition of types. The second section presents the relevant topological results from [Gro52], which are then applied in the third section to give a topological proof of the FTS, following the approach in [BY14]. In the final section, we establish an equivalent criterion for the stability of a formula and conclude with additional remarks and related results.

3.1 Basics of Stability Theory

Let x, y be disjoint finite multivariables. Let φ be an $\mathcal{L}(x, y)$-formula, T be an $\mathcal{L}$-theory, and $\boldsymbol{M}$ be a model of T. A *φ-instance* is a formula $\varphi(x, b)$ with $b \in M_y$. A *φ-formula* is an $\mathcal{L}_M$-formula which is equivalent (over $\boldsymbol{M}$) to one that is formed by a finite combination of non-quantifier connectives applied to φ-instances. A *φ-condition* is a condition of the form $\psi = 0$, where ψ is a φ-formula. A *φ-type* is a maximally satisfiable collection of φ-conditions.

Definition 3.1. The set of all φ-types is denoted by $\mathrm{S}_\varphi(M)$. For an element $a \in M_x$ we denote by $\mathrm{tp}_\varphi(a|M) \in \mathrm{S}_\varphi(M)$ the unique φ-type realized by a.

Let $\mathcal{F}$ be the set of all functions $\varphi(\dot{x}, b) : \mathrm{S}_x(M) \to [0, 1]$ for $b \in M_y$. It is easy to see that $\mathrm{S}_x(M)/\mathcal{F}$ is in a bijective correspondence with $\mathrm{S}_\varphi(M)$ by sending $[p] \in \mathrm{S}_x(M)/\mathcal{F}$, where $p \in \mathrm{S}_x(M)$ and $[p]$ is the class of p in $\mathrm{S}_x(M)/\mathcal{F}$, to the unique $q \in \mathrm{S}_\varphi(M)$ such that $q \subseteq p$ (we also denote $p|_\varphi = q$). We equip $\mathrm{S}_\varphi(M)$ with the quotient topology $\mathrm{S}_x(M)/\mathcal{F}$, which by Lemma 2.63 coincides with the initial topology on $\mathrm{S}_x(M)/\mathcal{F}$ generated by $\widetilde{\mathcal{F}}$ (we call this the *logic topology* on $\mathrm{S}_\varphi(M)$). By definition for any φ-formula ψ we have the continuous function $\dot{\psi}_{\mathrm{S}_\varphi(M)}$ (notation from Definition 2.15). By Proposition 1.4 we obtain that:

Proposition 3.2. *The set* $\{\dot{\psi}_{\mathrm{S}_\varphi(M)} : \psi$ *is a* φ*−formula*$\}$ *is dense in the set of continuous functions* $\mathrm{C}(\mathrm{S}_\varphi(M), [0, 1])$ *with respect to the uniform metric.*

Definition 3.3. We say that an M-definable x-predicate P is a *φ-predicate* if $\dot{P}$ factors through $\mathrm{S}_\varphi(M)$.

Proposition 3.2 leads to the following important corollary:

Corollary 3.4. *An M-definable x-predicate is a φ-predicate if and only if it is an uniform limit of a sequence $(\psi_n^{\boldsymbol{M}})$, where each ψ_n is a φ-formula.*

Definition 3.5. Let $\varepsilon > 0$. We say φ is *ε-stable* in $\boldsymbol{M}$ if there do not exist two infinite sequences $(a_n), (b_m)$ of elements of M_x, M_y such that for all $m < n$ we have $|\varphi(a_m, b_n) - \varphi(a_n, b_m)| \geq \varepsilon$. We call the formula φ *stable* in $\boldsymbol{M}$ if φ is ε-stable for every ε. Moreover, we say φ is *ε-stable* (in T) if φ is ε-stable in every model of T, and call φ *stable* if it is stable in every model of T.

The notion of definability readily translates from classical logic:

Definition 3.6. A type $p \in \mathrm{S}_\varphi(M)$ is called *definable* if the y-predicate $b \mapsto d_p\varphi(b) := \varphi(x,b)^p$ is M-definable. Here, $\varphi(x,b)^p$ is the unique real number such that the φ-condition $\varphi(x,b) = \varphi(x,b)^p$ belongs to p. We also define $\varphi^{op}(y,x) := \varphi(x,y)$.

Then we have the following important theorem:

Theorem 3.7. *[BYU10, Propositions 7.6,7.7] Suppose φ is stable. Then every φ-type $p \in \mathrm{S}_\varphi(M)$ is definable. Moreover $d_p\varphi$ is an M-definable φ^{op}-predicate.*

The proof provided in [BYU10] uses the monster model and Ramsey's Theorem, and is the continuous variant of the proof provided in [Pil96, Lemma 2.2]. Here we show a simpler and more natural proof, just using the definition of types, which also gives more of an insight into the definition of stability of a formula. Additionally, the proof we give also proves Theorem B.4 from [BYU10, Appendix B], where they prove that if φ is stable in $\boldsymbol{M}$ then every $p \in \mathrm{S}_\varphi(M)$ is definable.

The intuition behind our proof proceeds as follows. For $a \in M_x$ we have the continuous function $\varphi(\dot{a}, x) : \mathrm{S}_{\varphi^{op}}(M) \to [0,1]$, then in an attempt to generalize the element a we would consider $p(x) = \mathrm{tp}_\varphi(a|M)$ and try to instead build the function $p : \mathrm{S}_{\varphi^{op}}(M) \to [0,1]$ which sends q to $\varphi(a,y)^q$ for any $a \models p$, this is well-defined and continuous. Generalizing p, what if there is no realization of p? Then by trying to build the function $p : \mathrm{S}_{\varphi^{op}}(M) \to [0,1]$, to be consistent with the previous special cases, we could at least define it on the realized types $q \in \mathrm{S}_{\varphi^{op}}(M)$: for any $b \models q$, then $p(q) = \varphi(x,b)^p$, which is well-defined. The natural question is, when does there exist a continuous function $p : \mathrm{S}_{\varphi^{op}}(M) \to [0,1]$ such that $p(\mathrm{tp}_{\varphi^{op}}(b|M)) = \varphi(x,b)^p$. If this would be successful (for some special φ), then we could generally consider a map $\mathrm{S}_\varphi(M) \times \mathrm{S}_{\varphi^{op}}(M) \to [0,1], (p,q) \to p(q)$. Therefore we arrive at the following question: when does there exist a map (here we abuse notation for φ as a formula and as a map)

$$\varphi : \mathrm{S}_\varphi(M) \times \mathrm{S}_{\varphi^{op}}(M) \to [0,1],$$
$$\text{such that } \forall a \in M_x, \forall b \in M_y,\ \varphi(\mathrm{tp}_\varphi(a|M), \mathrm{tp}_{\varphi^{op}}(b|M)) = \varphi^{\boldsymbol{M}}(a,b),$$

such that for every $p \in \mathrm{S}_\varphi(M)$, $\varphi(p,\cdot)$ is continuous, and maybe even for every $q \in \mathrm{S}_{\varphi^{op}}(M)$, $\varphi(\cdot,q)$ is continuous. We continue to our proof.

Definition 3.8. For any $b \in M_y$ and $p \in \mathrm{S}_\varphi(M)$ we denote by b^p the value of $\varphi(x,b)^p$.

Lemma 3.9. *Suppose φ is stable, and let p be a φ-type and q be a φ^{op}-type. Then for any $\delta > 0$ exists $\varepsilon > 0$ and $a_0, \ldots, a_n \in M_x$ such that for any $b, b' \in M_y$ satisfying the $\mathcal{L}_M$-condition* $\max(|\varphi(a_0, y) - a_0^q|, \ldots, |\varphi(a_n, y) - a_n^q|) \leq \varepsilon$ *we have*

$$|b^p - (b')^p| \leq \delta.$$

Proof. Suppose that there exists $\delta > 0$ such that for all $\varepsilon > 0$ and $a_0, \ldots, a_n \in M_x$ exists $b, b' \in M_y$ satisfying $\max(|\varphi(a_0, y) - a_0^q|, \ldots, |\varphi(a_n, y) - a_n^q|) \leq \varepsilon$ and $|b^p - (b')^p| > \delta$.

Fix $\varepsilon = \delta/8$. For each n, we say that the elements $a_0, \ldots, a_n \in M_x$ and $b_0, \ldots, b_n \in M_y$ satisfy the property π if the following conditions hold:

1. for each $i \leq n$ we have $|a_i^q - b_i^p| \geq \delta/2$;
2. for each $i < j \leq n$ we have $|\varphi(a_i, b_j) - a_i^q| \leq \varepsilon$ and $|\varphi(a_j, b_i) - b_i^p| \leq \varepsilon$;

We will build by induction two $\mathbb{N}$-indexed sequences $(a_n) \subseteq M_x$ and $(b_n) \subseteq M_y$ such that for every n the sequences $a_0, \ldots, a_n$ and $b_0, \ldots, b_n$ satisfy property π. If such a sequence exists, observe that for every $m < n$ we have

$$|\varphi(a_m, b_n) - \varphi(a_n, b_m)| \geq |a_m^q - b_m^p| - 2\varepsilon \geq \delta/4 > 0,$$

a contradiction to the stability of φ.

For the base step, let $a_0 \in M_x$ be an arbitrary element. Then there exist two elements $b_0, b_0' \in M_y$ such that $|b_0^p - (b_0')^p| > \delta$. Then $\delta < |a_0^q - b_0^p| + |a_0^q - (b_0')^p|$, hence we can suppose $|a_0^q - b_0^p| \geq \delta/2$ (otherwise we pick b_0 to be b_0'). Suppose, $a_0, \ldots, a_n$ and $b_0, \ldots, b_n$ satisfy property π. Since p is φ-type there exists an element $a_{n+1} \in M_x$ such that $|\varphi(a_{n+1}, b_m) - b_m^p| \leq \varepsilon$ for every $m \leq n$. By our assumption there exists $b_{n+1}, b_{n+1}' \in M_y$ both satisfying $\max(|\varphi(a_1, y) - a_1^q|, \ldots, |\varphi(a_n, y) - a_n^q|) \leq \varepsilon$ and $|b_{n+1}^p - (b_{n+1}')^p| > \delta$. Finally by similar reasoning as at the base step we can suppose that $|a_{n+1}^p - b_{n+1}^q| \geq \delta/2$, completing the induction step. $\square$

Now, suppose φ is stable, fix $p \in \mathrm{S}_\varphi(M)$ and we will build a function $p : \mathrm{S}_{\varphi^{op}}(M) \to [0, 1]$. Let $q \in \mathrm{S}_{\varphi^{op}}(M)$ and let an open set $q \in O \subseteq \mathrm{S}_{\varphi^{op}}(M)$. Consider

$$I_O = cl(\{b^p : \mathrm{tp}_{\varphi^{op}}(b|M) \in O\}) \subseteq [0, 1],$$

where $cl(\cdot)$ is the closure of a subset of real numbers. Clearly if $O \subseteq U$ then $I_O \subseteq I_U$. Therefore if for any $\delta > 0$ there exists an open neighborhood O of q such that the diameter of I_O is $\leq \delta$, then there exists an unique element

$$p(q) := \bigcap\{I_O : O \text{ an open neighborhood of } q \text{ in } \mathrm{S}_{\varphi^{op}}(M)\}. \tag{3.1}$$

The existence of this element is guaranteed by Lemma 3.9 which says that there exists an open neighborhood $O \subseteq \mathrm{S}_{\varphi^{op}}(M)$ of q such that for any b, $b' \in M_y$ with $\mathrm{tp}_{\varphi^{op}}(b|M)$, $\mathrm{tp}_{\varphi^{op}}(b'|M) \in O$ we have $|b^p - (b')^p| \leq \delta$. The continuity of this function is obvious as, for example, given $q \in \mathrm{S}_{\varphi^{op}}(M)$ let O be an open set containing q such that diameter of I_O is $\leq \delta$. Then $d(p(q), p(O)) \leq \delta$. It is also easy to see that for any $b \in M_y$ we have $p(\mathrm{tp}_{\varphi^{op}}(b|M)) = b^p = (\varphi(x,b))^p$. And since the set

$$\{\mathrm{tp}_{\varphi^{op}}(b|M) : b \in M_y\}$$

is dense in $\mathrm{S}_{\varphi^{op}}(M)$ we obtain the following:

Corollary 3.10. *Suppose φ is stable. Then there is an unique continuous function $p : \mathrm{S}_{\varphi^{op}}(M) \to [0,1]$ such that for every $b \in M_y$ we have $p(\mathrm{tp}_{\varphi^{op}}(b|M)) = (\varphi(x,b))^p$.*

Proof of Theorem 3.7. Let p be a φ-type. Then $p : \mathrm{S}_{\varphi^{op}}(M) \to [0,1]$ is continuous hence there exists a M-definable φ^{op}-predicate P such that $p = \dot{P}$. This clearly implies that for any $b \in M_y$ we have $p(\mathrm{tp}_{\varphi^{op}}(b|M)) = \dot{P}(\mathrm{tp}_{\varphi^{op}}(b|M))$ and respectively $\varphi(x,b)^p = P(b)$. □

Observe that in the proof of Lema 3.9 we only used the fact that φ is stable in $\boldsymbol{M}$, therefore we get:

Theorem 3.11. *[BYU10, Theorem B.4] Suppose φ is stable in $\boldsymbol{M}$. Then every φ-type in $\mathrm{S}_\varphi(M)$ is definable. Moreover $d_p\varphi$ is an M-definable φ^{op}-predicate.*

We also state the following slight generalization of Lemma 3.9, from which Lemma 3.9 follows as an immediate corollary.

Lemma 3.12. *Suppose φ is stable in $\boldsymbol{M}$. For any $p \in \mathrm{S}_\varphi(M)$, $q \in \mathrm{S}_{\varphi^{op}}(M)$ and for any $\delta > 0$ there exists $\varepsilon > 0$ and $a_1, \ldots, a_n \in M_x, b_1 \ldots, b_n \in M_y$ such that for any $a \in M_x$ and $b \in M_y$ satisfying*

$$a \models \max(|\varphi(x,b_1) - b_1^p|, \ldots, |\varphi(x,b_n) - b_n^p|) \leq \varepsilon \tag{3.2}$$
$$b \models \max(|\varphi(a_1,y) - a_1^q|, \ldots, |\varphi(a_n,y) - a_n^q|) \leq \varepsilon \tag{3.3}$$

we have

$$|a^q - b^p| \leq \delta. \tag{3.4}$$

Proof. Suppose that there exists $\delta > 0$ such that for all $\varepsilon > 0$ and sequences $a_0, \ldots, a_n \in M_x$, $b_0, \ldots, b_n \in M_y$ there exists $a \in M_x, b \in M_y$ satisfying conditions (3.2) and (3.3) but $|a^q - b^p| > \delta$. Fix $\varepsilon = \delta/4$.
We say that the sequence $a_0, \ldots, a_n, b_0, \ldots, b_n$ satisfies the property π if

(i) for all $i < j \leq n$ we have $|\varphi^{\boldsymbol{M}}(a_i,b_j) - a_i^q| \leq \varepsilon$ and $|\varphi^{\boldsymbol{M}}(a_j,b_i) - b_i^p| \leq \varepsilon$;

(ii) for each $m \leq n$ we have $|a_m^q - b_m^p| > \delta$.

We will build by induction two sequences $(a_n) \subseteq M_x, (b_n) \subseteq M_y$ such that for each n the sequence $a_0, \ldots, a_n, b_0, \ldots, b_n$ satisfies the property π. In such a case observe that for every $m < n$ we obtain

$$|\varphi^{\boldsymbol{M}}(a_m,b_n) - \varphi^{\boldsymbol{M}}(a_n,b_m)| \geq |a_m^q - b_m^p| - 2\varepsilon \geq \delta - 2\varepsilon = \delta/2 > 0,$$

which contradicts the stability of φ.

For the base step, we can obviously pick $a_0 \in M_x, b_0 \in M_y$ such that $|a_0^q - b_0^p| \geq \delta$. Suppose that we have built the sequence $a_0, \ldots, a_n, b_0, \ldots, b_n$ which satisfies property π. Then, by our initial assumption at the start of the proof, we can pick $a_{n+1} \in M_x$ and $b_{n+1} \in M_y$ such that

$$\begin{aligned} a_{n+1} &\models \max(|\varphi(x, b_0) - b_0^p|, \ldots, |\varphi(x, b_n) - b_n^p|) \leq \varepsilon \\ b_{n+1} &\models \max(|\varphi(a_0, y) - a_0^q|, \ldots, |\varphi(a_n, y) - a_n^q|) \leq \varepsilon \end{aligned}$$

and $|a_{n+1} - b_{n+1}| > \delta$. Finally, we observe that the sequence $a_0, \ldots, a_{n+1}, b_0, \ldots, b_{n+1}$ satisfies property π. ☐

We observe that for φ not necessarily stable, if $a, a' \models p$ and $b, b' \models q$ then $\varphi^{\boldsymbol{M}}(a, b) = \varphi^{\boldsymbol{M}}(a', b')$. Thus, when attempting to define a function $\varphi : \mathrm{S}_x(M) \times \mathrm{S}_y(M) \to [0, 1]$ we can at least define φ on the pairs of realized types in $\boldsymbol{M}$, which form a dense set in the space $\mathrm{S}_x(M) \times \mathrm{S}_y(M)$. Moreover, if φ is stable, then by Lemma 3.12, we can consider the function

$$\varphi : \mathrm{S}_x(M) \times \mathrm{S}_y(M) \to [0, 1], (p, q) \to p(q) = q(p), \tag{3.5}$$

where $\varphi(\mathrm{tp}_{\varphi}(a|M)), \mathrm{tp}_{\varphi^{op}}(b|M)) = \varphi^{\boldsymbol{M}}(a, b)$, and for each $\varphi(p, \cdot), \varphi(\cdot, q)$ are continuous for each $p \in \mathrm{S}_{\varphi}(M), q \in \mathrm{S}_{\varphi^{op}}(M)$. The identity $p(q) = q(p)$, known as the symmetry property [BYI25, Theorem 2.7(*ii*)], follows from inequality (3.4).

Remark 3.13. *The space* $\mathrm{S}_{\varphi}(M)$ *can be equipped with a metric defined by* $d(p, p') = \sup_b |b^p - b^{p'}|$ *(if* φ *is stable, then this represents the uniform metric between the maps* $p : \mathrm{S}_{\varphi^{op}}(M) \to [0, 1]$ *and* $p' : \mathrm{S}_{\varphi^{op}}(M) \to [0, 1]$*). The metric topology generated by* d *is finer than the logic topology on* $\mathrm{S}_{\varphi}(M)$*. If* φ *is stable, then we can build the function*

$$\Phi : \mathrm{S}_{\varphi}(M) \to \mathrm{C}(\mathrm{S}_{\varphi^{op}}(M), [0, 1]), \ p \to d_{\varphi} p,$$

which is an isometry on its image, with respect to the d*-metric on* $\mathrm{S}_{\varphi}(M)$ *and the uniform metric on* $\mathrm{C}(\mathrm{S}_{\varphi^{op}}(M), [0, 1])$*. In the theorems below,* $\|\mathrm{S}_{\varphi}(M)\|$ *is with respect to the metric topology.*

Finally, we obtain:

Theorem 3.14 (The Fundamental Theorem of Stability)**.** *[BYU10, Proposition 7.7] The following are equivalent:*

(i) φ *is stable;*

(ii) for every $\boldsymbol{M} \models T$*, every* $p \in \mathrm{S}_{\varphi}(M)$ *is definable;*

(iii) for every infinite $\kappa \geq |\mathcal{L}|$ *and* $\boldsymbol{M} \models T$*, with* $\|\boldsymbol{M}\| \leq \kappa$*, we have* $\|\mathrm{S}_{\varphi}(M)\| \leq \kappa$*;*

(iv) there exists $\kappa \geq |\mathcal{L}|$ *such that for every* $\boldsymbol{M} \models T$*, with* $\|\boldsymbol{M}\| \leq \kappa$*, we have* $\|\mathrm{S}_{\varphi}(M)\| \leq \kappa$*;*

(v) there exists $\kappa \geq |\mathcal{L}|$, with $\kappa^{\aleph_0} = \kappa$, such that for every $\boldsymbol{M} \models T$, with $|\boldsymbol{M}| \leq \kappa$, we have $|\mathrm{S}_\varphi(M)| \leq \kappa$.

Proof. The implication $(i) \Rightarrow (ii)$ follows from Theorem 3.7. Suppose (ii) and let $\boldsymbol{M} \models T$, with $\|\boldsymbol{M}\| \leq \kappa$ and $\kappa \geq |\mathcal{L}|$. Then for a dense subset $D \subseteq M$, by using that formulas are uniformly continuous and Proposition 3.2 it is obvious that $\|\mathrm{C}(\mathrm{S}_\varphi(M), [0,1])\| \leq |\mathcal{L}| \cdot |D|$, and then

$$\|\mathrm{S}_\varphi(M)\| = \|\mathrm{im}(\Phi)(\mathrm{S}_\varphi(M))\| \leq \|\mathrm{C}(\mathrm{S}_\varphi(M), [0,1])\| \cdot \aleph_0 \leq |\mathcal{L}| \cdot \|\boldsymbol{M}\| \cdot \aleph_0 = \kappa,$$

the first $\leq$ is a simple property between the size (as in the smallest cardinality of a dense set) of subspace of a metric space in relation to the size of the metric space, thereby proving (iii). The implication $(iii) \implies (iv)$ is obvious, while the implication $(iv) \implies (i)$ is standard and can be seen in the proof of [BYU10, Proposition 7.7]. The implication $(iii) \implies (v)$ results from choosing an infinite cardinal $\kappa^{\aleph_0} = \kappa$. Lastly the implication $(v) \implies (iv)$ is trivial since $|\boldsymbol{M}| \leq \|\boldsymbol{M}\|^{\aleph_0}$. □

Definition 3.15. A theory T is called *stable* if every partitioned $\mathcal{L}$-formula $\varphi(x, y)$ is stable in T. A type $p \in \mathrm{S}_x(M)$ is called *definable* if for every finite multivariable z and ψ an $\mathcal{L}(x, z)$-formula, $p|_\psi$ is definable.

Definition 3.16. Let λ be a cardinal. We say that T is *λ-stable* if for every $A \subseteq \boldsymbol{M} \models T$ with $|A| \leq \lambda$, and every finite multivariable x , we have $\|\mathrm{S}_x(A)\| \leq \lambda$. Here, and in the following results, $\|\mathrm{S}_x(A)\|$ is with respect to the metric topology, as in Theorem 1.80.

Finally, we arrive at the general theorem:

Theorem 3.17. *[BYU10, Theorem 8.5] The following are equivalent:*

(i) T is stable;

(ii) For all $\boldsymbol{M} \models T$, and every finite x, all the types in $\mathrm{S}_x(M)$ are definable;

(iii) T is λ-stable for some $\lambda \geq |\mathcal{L}|$.

Proof. All the implications use Theorem 3.14 and counting arguments, a proof can be seen at [BYU10, Theorem 8.5]. □

Definition 3.18. Let λ be a cardinal. We say that T is *λ-stable with respect to the discrete metric* if for every $A \subseteq \boldsymbol{M} \models T$ with $|A| \leq \lambda$, and every finite multivariable x, we have $|\mathrm{S}_x(A)| \leq \lambda$. We say T is *stable with respect to the discrete metric* if it is λ-stable with respect to the discrete metric, for some λ.

Furthermore, the two notions of stability, stable and stable with respect to the discrete metric, are equivalent, a fact proved in Theorem 14.6 of [BBHU08].

We conclude this section with a theorem demonstrating that, in the definition of stability, it suffices to consider x as a singleton.

Proposition 3.19. *Let v be a single variable, and suppose for every $A \subseteq \boldsymbol{M} \models T$ with $|A| \leq \lambda$ we have $\|\mathrm{S}_v(A)\| \leq \lambda$, then T is λ-stable.*

Proof. We prove this by induction on n, working in a monster model $\mathbb{M} \models T$. Suppose, for every $A \subseteq \mathbb{M}$ with $|A| \leq \lambda$, and $|x| \leq n$ we have $\|\mathrm{S}_x(A)\| \leq \lambda$. Let $x = (y, v)$ be a multivariable of length $n+1$. Let $D \subseteq \mathrm{S}_v(A)$ be a dense set of cardinality $\leq \lambda$. We define

$$\mathcal{D} = \{p \in \mathrm{S}_{yv}(A) : p|_v \in \mathrm{S}_v(A)\}.$$

Then $\mathcal{D}$ is dense in $\mathrm{S}_{yv}(A)$. Indeed, let $p \in \mathrm{S}_{yv}(A)$, let $\varepsilon > 0$, and let $(a, b) \in \mathbb{M}_{yv}$ be a realization of p. Let $r \in D$ and $b' \in \mathbb{M}$ a realization of r such that $d(p|v, r) = d^{\mathbb{M}}(b, b') \leq \varepsilon$. Then for $q = \mathrm{tp}(a, b') \in \mathcal{D}$ we have that $d(p, q) \leq d^{\mathbb{M}}((a, b), (a, b')) \leq \varepsilon$.

Next, for each $r \in D$ we choose $b_r \in \mathbb{M}$ such that $b_r \models r$, and let $\mathcal{D}_r = \{p \in \mathcal{D} : r \subseteq p\}$. Then, we can build a bijection $f_r : \mathcal{D}_r \to \mathrm{S}_y(A \cup \{b_r\}), q \to q(y, b_r)$. Note that this map is a contraction, that is $d(q, q') \leq d(q(y, b_r), q'(y, b_r))$. Hence, since $\mathrm{S}_y(A \cup \{b_r\})$ has a dense set E_r of cardinality $\leq \lambda$ then $f_r^{-1}(E_r) = D_r$ is a dense set in $\mathcal{D}_r$ of cardinality $\leq \lambda$.

Finally, since $\bigcup_r \mathcal{D}_r = \mathcal{D}$, it follows that $\bigcup_r D_r$ is dense in $\mathcal{D}$, and hence dense in $\mathrm{S}_{yv}(A)$. Moreover, since $|D| \leq \lambda$ and each $|D_r| \leq \lambda$, we obtain $|\bigcup_r D_r| \leq \lambda \cdot \lambda = \lambda$. $\square$

3.2 Convergence in Banach Spaces

In this section, we outline the key topological notions that will be used throughout, and present the relevant results from [Gro52].

Let X be a topological space. We denote by $\mathrm{C}(X)$ the space $\mathrm{C}(X, \mathbb{C})$, and by $\mathrm{C}_b(X)$ the subset of bounded functions from $\mathrm{C}(X)$. We equip $\mathrm{C}_b(X)$ with the supremum norm, making it a Banach space. We denote by $\mathrm{C}_s(X)$ the set $\mathrm{C}(X, \mathbb{C})$ with the topology of pointwise convergence, i.e. the topology inherited from the product topology on $\mathbb{C}^X$.

Definition 3.20. A point $x \in X$ is called a *cluster point* of a sequence (x_n) in X if for every neighborhood V of x, there are infinitely many n such that $x_n \in V$. For a subset A of X, we call $x \in X$ a *limit point* of A if every neighborhood of x contains a point of A other than x itself.

Observe that if X is Hausdorff, then $x \in X$ is a limit point of $A \subseteq X$ if and only if x is an ω-accumulation point of A (i.e. every neighborhood of x contains infinitely many points from A).

Definition 3.21. A subset $A \subseteq X$ is said to be:

(i) *relatively compact* if its closure in X is compact;

(ii) *sequentially compact* (respectively, *relatively sequentially compact*) if every sequence from A has a convergent subsequence in A (respectively, in X) ;

(iii) *cluster-sequentially compact* (respectively, *relatively cluster-sequentially compact*) if every sequence from A has a cluster point in A (respectively, in X).

If A is (relatively) compact or (relatively) sequentially compact then A is (relatively) cluster-sequentially compact. Observe that X is cluster-sequentially compact if and only if every infinite subset of X has an ω-accumulation point. If X is a metric space, then for any subset of X the notions of relatively compact, relatively sequentially compact and relatively cluster-sequentially compact are all equivalent. In this subsection we will prove that this equivalence also holds for more complicated spaces.

Remark 3.22. *Cluster-sequential compactness of X is equivalent with X being countably compact (i.e. every countable open cover of X has a finite subcover) (see [SS95, p. 19]). Hence, for a subset A of X, if the closure of A is countably compact, then A is relatively cluster-sequentially compact. On the other hand, A being relatively cluster-sequentially compact does not imply the closure of A is cluster-sequentially compact, equivalently countably compact.*

Remark 3.23. *Let A be a subset of $\mathrm{C}_s(X)$. Then A is relatively compact in $\mathrm{C}_s(X)$ if and only if:*

(i) for every $x \in X$ the set $\{f(x) : f \in A\}$ is relatively compact in $\mathbb{C}$ and,

(ii) the closure of A in $\mathrm{C}_s(X)$ coincides with the closure of A in $\mathbb{C}^X$.

For the $\Rightarrow$ direction: it is easy to see that (i) holds, and since $\mathbb{C}^X$ is Hausdorff we obtain that the closure of A in $\mathrm{C}_s(X)$, which is compact, coincides with the closure of A in $\mathbb{C}^X$. For the $\Leftarrow$ direction, using Tychonoff's Theorem we obtain that if A satisfies properties (i) and (ii) then A is relatively compact in $\mathrm{C}_s(X)$.

Remark 3.24. *Suppose $x \in X$ is a cluster point of some sequence $(x_n) \subseteq X$, and let $f \in \mathrm{C}(X)$ be given. Then, by continuity, $f(x)$ is also a cluster point of the sequence $(f(x_n))$ in $\mathbb{C}$. Hence, if for some open set $V \subseteq \mathbb{C}$ there are only finitely many n such that $f(x_n) \in V$, then $f(x) \notin V$.*

Similarly if f is a cluster point of some sequence $(f_n) \subseteq C_s(X)$, then for any $x \in X$, $f(x)$ is a cluster point of the complex-valued sequence $(f_n(x))$. Moreover, in $\mathbb{C}$, a cluster point of a sequence is nothing but a limit of a subsequence.

The following is a proof of a variant of Theorem 2 from [Gro52], which suffices for our purposes.

Theorem 3.25. *[Gro52, Theorem 2] Suppose X is cluster-sequentially compact (equivalently countably compact), A a subset of $C_s(X)$, and X_0 a dense subset of X. Then the following are equivalent:*

(i) A is relatively compact;

(ii) A is relatively cluster-sequentially compact;

(iii) for every sequence $(f_n) \subseteq A$ and every sequence $(x_m) \subseteq X_0$, there exists a double cluster point $c \in \mathbb{C}$ of the double sequence $(f_n(x_m))$ (meaning that every neighborhood of c meets infinitely many rows and infinitely many columns of the double

sequence, each in infinitely many terms). Furthermore, for every $x \in X \setminus X_0$, the set $\{f(x) : f \in A\}$ is relatively cluster-sequentially compact in $\mathbb{C}$ (equivalently bounded).

Moreover the implication (iii) $\implies$ (i) holds for any topological space X.

Proof. The implication $(i) \to (ii)$ is straightforward. For the direction $(ii) \to (iii)$, suppose (ii), and let sequences $(f_n) \subseteq A$ and $(x_m) \subseteq X_0$ be given. Let f be a cluster point of the sequence (f_n) and $x \in X$ a cluster point of (x_m). It is obvious that for every $y \in X$ the set $\{f(y) : f \in A\}$ is relatively cluster-sequentially compact (because of the topology on $\mathrm{C}_s(X, \mathbb{C})$). In the table below an arrow indicates that the indicated element is a cluster point of the respective sequence.

By Remark 3.24 we obtain: for each n, $f_n(x)$ is a cluster point of the sequence $(f_n(x_m))_m$; for each m, $f(x_m)$ is a cluster point of $(f_n(x_m))_n$; and finally $f(x)$ is a cluster point of both $(f_n(x))_n$ and $(f(x_m))_m$.

	x_0	x_1	x_2	$\to$	x
f_0	$f_0(x_0)$	$f_0(x_1)$	$f_0(x_2)$	$\to$	$f_0(x)$
f_1	$f_1(x_0)$	$f_1(x_1)$	$f_1(x_2)$	$\to$	$f_1(x)$
f_2	$f_2(x_0)$	$f_2(x_1)$	$f_2(x_2)$	$\to$	$f_2(x)$
$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$		$\downarrow$
f	$f(x_0)$	$f(x_1)$	$f(x_2)$	$\to$	$f(x)$

All it remains is to observe that $c := f(x)$ is a double cluster point of $(f_n(x_m))$.

Suppose (iii). For every $x \in X \setminus X_0$ the set $\{f(x) : f \in A\}$ is relatively cluster-sequentially compact, which in $\mathbb{C}$ is equivalent to relatively compact. While for $x \in X_0$ by choosing the constant sequence $(x)_n$ in condition (iii) we obtain that $\{f(x) : f \in A\}$ is also relatively compact in $\mathbb{C}$. Hence by Remark 3.23, we are left with proving that the closure of A in $\mathrm{C}_s(X)$ is the same as the closure of A in $\mathbb{C}^X$. For this to be true we need to prove that any element from the closure of A in $\mathbb{C}^X$ is continuous.

Let $f : X \to \mathbb{C}$ be an element from the closure of A in $\mathbb{C}^X$. It is straightforward to prove that if for any $x \in X$ and any $\varepsilon > 0$ there exists a neighborhood V of x such that the diameter of $\{f(x)\} \cup f(V \cap X_0)$ is $\leq \varepsilon$, then f is continuous at x. Towards a contradiction suppose $x \in X$ and $\varepsilon > 0$ are such that for every neighborhood V of x there exists an element $x' \in V \cap X_0$ with $|f(x) - f(x')| \geq \varepsilon/2$. We build recursively two sequences $(f_n)_{n \geq 1} \subseteq A$ and $(x_m)_{m \geq 1} \subseteq X_0$ which satisfy the following three conditions:

(1) for each $n \geq 1$ and $1 \leq m < n$ we have $|f(x) - f_n(x)| \leq \frac{1}{n}$ and $|f(x_m) - f_n(x_m)| \leq \frac{1}{n}$;

(2) for each $n \geq 1$ and $1 \leq m \leq n$ we have $|f_m(x) - f_m(x_n)| \leq \frac{1}{n}$;

(3) for each $m \geq 1$ we have $|f(x) - f(x_m)| \geq \frac{\varepsilon}{2}$.

The recursion step is: suppose $f_1, \ldots, f_n$ and $x_1, \ldots, x_n$ have been chosen, then we pick $f_{n+1} \in A$ such that f_{n+1} is $\frac{1}{n+1}$ close to f on the domain $\{x, x_1, \ldots, x_n\}$ and then we pick

$$x_{n+1} \in f_1^{-1}(B(f_1(x), 1/(n+1))) \cap \cdots \cap f_{n+1}^{-1}(B(f_{n+1}(x), 1/(n+1))) \cap X_0,$$

such that $|f(x) - f(x_{n+1})| \geq \frac{\varepsilon}{2}$ (the notation $B(c, \delta)$ is for the open ball in $\mathbb{C}$ with center c and radius δ). The arrow in the table below indicates the limit of the respective sequence.

By (iii), the double sequence $(f_n(x_m))_{n\geq 1,m\geq 1}$ has a double cluster point $c \in \mathbb{C}$. Then c is a cluster point of $(f(x_m))$ by (1) and a cluster point of $(f_n(x))$ by (2). Hence $c = f(x)$ since $\lim_n f_n(x) = f(x)$.

$$
\begin{array}{c:c|ccccc:c}
 & x & x_1 & x_2 & x_3 & \cdots & & \\
\hline
f_1 & f_1(x) & f_1(x_1) & f_1(x_2) & f_1(x_3) & \to & & f_1(x) \\
f_2 & f_2(x) & f_2(x_1) & f_2(x_2) & f_2(x_3) & \to & & f_2(x) \\
f_3 & f_3(x) & f_3(x_1) & f_3(x_2) & f_3(x_3) & \to & & f_3(x) \\
\vdots & \downarrow & \downarrow & \downarrow & \downarrow & & & \downarrow \\
\hdashline
 & f(x) & f(x_1) & f(x_2) & f(x_3) & \not\to & & f(x)
\end{array}
$$

This is not possible as by condition (3) c cannot be a cluster point of $(f(x_m))_{m\geq 1}$. $\square$

It is interesting to point out that for the double sequence $(f_n(x_m))_{n\geq 1,m\geq 1}$ which is built by recursion, the end result is: (f_n) converges pointwise on $\{x, x_1, x_2, \dots\}$ and the sequence (x_m) converges to x in the coarsest topology on X making the functions (f_n) continuous (observe that this topology is first countable so the only kind of convergence (using nets) which matters is the sequence indexed by $\mathbb{N}$). This can be done because in the process of recursion we only need to use open sets which witness the fact that each f_n is continuous, and we can force, when we choose recursively the elements x_m, for the sequence (x_m) to converge to x. In the proof we only use the fact that $\lim_n f_n(x)$ is a cluster point of the sequence $(\lim_n f_n(x_m))_m$. But then we could even restrict the index set of the columns to $(m_k)_k$ so that the $\lim_n f_n(x) = \lim_k \lim_n f_n(x_{m_k})$. Summarizing, the only kind of double sequence required for the proof is: a sequence (f_n) and a sequence (x_m) which converges to x (in the coarsest topology on X making the functions f_n continuous), such that (f_n) converges pointwise on $\{x, x_1, x_2, \dots\}$ and then if $\lim_m \lim_n f_n(x_m)$ exists then it must be equal to $\lim_n f_n(x)$. Which is actually just interchangeability of limits

$$\lim_m \lim_n f_n(x_m) = \lim_n f_n(x) = \lim_n \lim_m f_n(x_m).$$

Remark 3.26. *In [Gro52, Theorem 2], Theorem 3.25 is proved in a more general case; instead of considering complex-valued functions we consider F-valued functions, for F a Tychonoff space (which by definition is completely regular and hausdorff), and the statement of the theorem remains unchanged. The proof follows the same lines plus uses the fact that a Tychonoff space has the initial topology with respect to a set of complex-valued functions (Tychonoff spaces embed in some cube $[0,1]^I$).*

Remark 3.27. *Let $(x_{n,m})$ be a double sequence in a metric space such that $\{x_{n,m} : m, n \geq 0\}$ is relatively compact. Then it is fairly easy using diagonal arguments to extract a double subsequence $(x_{n_i,m_j})_{i,j\in\mathbb{N}}$ such that both $\lim_i \lim_j x_{n_i,m_j}$ and $\lim_j \lim_i x_{n_i,m_j}$ exist. This is done by restricting m to a subsequence of indices $I_1 \subseteq \mathbb{N}$ such that the first row converges, then furthermore restricting m to a subsequence of indices $I_2 \subseteq I_1$ such that the second row converges, this procedure continues for all rows, then we apply a diagonal argument and restrict the values of m to be from this diagonal; then for this new double sequence do the same procedure for the columns; we are left with a double sequence which converges in any row and column; finally all it remains is to restrict the indices of the row and the column, such that the sequences of the limits of the rows and* *limits of the columns are both convergent. Therefore a relatively compact double sequence* $(x_{n,m})$ *has no cluster point implies that for any double subsequence* $(x_{n_i,m_j})_{i,j\in\mathbb{N}}$*, if both* $\lim_i \lim_j x_{n_i,m_j}$ *and* $\lim_j \lim_i x_{n_i,m_j}$ *exist, then they are distinct.*

Moreover if the double sequence $(x_{n,m})$ *has a cluster point and both* $\lim_n \lim_m x_{n,m}$ *and* $\lim_m \lim_n x_{n,m}$ *exist, then* $\lim_n \lim_m x_{n,m} = \lim_m \lim_n x_{n,m}$ *and* $\lim_n \lim_m x_{n,m}$ *is the cluster point.*

By applying the preceding remark in conjunction with Theorem 3.25, we derive the following result:

Corollary 3.28. *[Gro52, Corollary 3] Using the notations of Theorem 3.25, suppose the set* $\{f(x) : f \in A, x \in X\}$ *is bounded. Then* A *is relatively compact in* $C_s(X)$ *if and only if for any two sequences* $(f_n) \subseteq A$ *and* $(x_m) \subseteq X_0$ *we have*

$$\lim_m \lim_n f_n(x_m) = \lim_n \lim_m f_n(x_m)$$

provided both limits exist. Furthermore, this condition is sufficient to ensure that A *is relatively compact, even if* X *is not assumed to be cluster-sequentially compact.*

Observe that in condition (iii) of Theorem 3.25, the topology on X matters only up to the point of making the functions from A continuous. Therefore, in the case $X_0 = X$ and condition (iii) is true for some topology on X, then it is also true for X equipped with the smallest topology τ making all functions from A continuous; and then all elements from the closure of A in $\mathbb{C}^X$ are continuous with respect to τ.

Proposition 3.29. *[Gro52, Theorem 4] Suppose* X *is a compact space and let* A *be a relatively cluster-sequentially compact subset of* $\mathrm{C}_s(X)$*. Then* A *is relatively sequentially compact.*

Proof. Let (f_n) be a sequence in A. We need to show (f_n) has a convergent subsequence in $\mathrm{C}_s(X)$. On X let τ be the smallest topology making all the functions $\{f_n : n \geq 0\}$ continuous, denote this topological space by $\tilde{X}$. Then, by Theorem 3.25, $\{f_n : n \geq 0\}$ is relatively compact (and relatively cluster-sequentially compact) in $\mathrm{C}_s(\tilde{X})$, and the closure of $\{f_n : n \geq 0\}$ in $\mathbb{C}^X$ is contained in $\mathrm{C}_s(\tilde{X})$.

Suppose there exists a subsequence (f_{n_k}) and a dense set X_0 in $\tilde{X}$, such that for all $x \in X_0$, $\lim_k f_{n_k}(x)$ exists. Then (f_{n_k}) has a unique cluster point f: a cluster point exists since A is relatively cluster-sequentially compact, and if g is another cluster point, since $f = g$ on X_0, then $f = g$ on X as well. Morover for any $x \in X\backslash X_0$ we have $\lim_k f_{n_k}(x) = f(x)$, because otherwise we can pass to a subsequence $(f_{n_{k_l}})$ with a cluster point g where $g(x) \neq f(x)$, and g would also be a cluster point of (f_{n_k}), a contradiction. Summarizing, we obtained a function $f \in \mathrm{C}_s(\tilde{X}) \subseteq \mathrm{C}_s(X)$ as the pointwise limit of (f_{n_k}). Therefore is enough to find a dense set X_0 and a subsequence f_{n_k} converging pointwise on X_0.

The topology τ is also the smallest topology on X for which the function

$$\iota : X \to \prod_n f_n(X), \text{ defined by } x \mapsto (f_n(x))_n,$$

is continuous. Observe that since $f_n(X)$ is compact, from basic topology we know $\prod_n f_n(X)$ is metrizable and also compact by Tychonoff's Theorem. Consequently $\iota(X)$ is a compact metric space, hence separable, and τ must be a compact separable pseudometric space. What is important is that we can pick a dense countable set X_0 of $\tilde{X}$.

Finally, we enumerate $X_0 = \{x_0, x_1, \dots\}$ and consider a cluster point $f \in \mathrm{C}_s(\tilde{X})$ of (f_n). Then recursively we can construct a subsequence (f_{n_k}) such that for all $m \leq k$ we have $|f(x_m) - f_{n_k}(x_m)| \leq \frac{1}{m+1}$. As a consequence, for every $x \in X_0$, the limit $\lim_k f_{n_k}(x)$ exists. □

From Theorem 3.25 and Proposition 3.29, we obtain:

Theorem 3.30. *Suppose X is a compact space and let A be a subset of $\mathrm{C}_s(X)$. Then the following are equivalent:*

(i) A is relatively compact;

(ii) A is relatively sequentially compact;

(iii) A is relatively cluster-sequentially compact.

Now, we need to introduce some basic notions and theorems from the theory of Banach spaces. Let E be a Banach space. By E^* we denote the Banach space of all continuous linear functionals of E. The *weak topology* on E is the coarsest topology making all the functions from E^* continuous. It is easy to see that the weak topology is Hausdorff and coarser than the norm topology on E.

Definition 3.31. Let A be a subset of E. By the *weak closure* of A we mean the closure of A in the weak topology of E. We say that A is *weakly compact* if A is compact in the weak topology, and we say that A is *weakly relatively compact* if the weak closure of A is weakly compact. In the obvious way we define the notions of *weak convergence*, *weak limit*, and a *weak cluster point* of a sequence in E. Similarly we define for A to be *weakly (relatively) sequentially compact* and *weakly (relatively) cluster-sequentially compact* (weakly means that these notions are with respect to the weakly topology).

Remark 3.32. *Immediately from the definitions we see that a sequence (x_n) from E converges weakly to $x \in E$ if and only if for all continuous functionals $T \in E^*$ we have $T(x_n) \to T(x)$.*

The following is a deep theorem concerning the weak topology of Banach spaces.

Theorem 3.33 (Eberlein–Šmulian Theorem)**.** *[Con19, p.163] Let $A \subseteq E$. Then the following are equivalent:*

(i) A is weakly relatively compact;

(ii) A is weakly relatively sequentially compact;

(iii) A is weakly relatively cluster-sequentially compact.

Using Theorem 3.25 (with F a completely regular space instead of $\mathbb{C}$) Grothendieck provides in Proposition 2 of his article [Gro52], a generalization of the Eberlein–Šmulian Theorem to the case of E a locally convex space.

Suppose X is a compact space. Then $\mathrm{C}(X)$ is a Banach space with the supremum norm. By the Riesz–Markov–Kakutani representation theorem, all the continuous linear functionals on $\mathrm{C}(X)$ are given by Radon measures on X. Hence a sequence of functions $(f_n) \in \mathrm{C}(X)$ converges weakly to $f \in \mathrm{C}(X)$ if and only if for all Radon measures μ on X we have $\int f_n\mu \to \int f\mu$.

Fact 3.34. *Suppose X is compact. A sequence (f_n) in $\mathrm{C}(X)$ converges weakly to $f \in \mathrm{C}(X)$ if and only if (f_n) converges pointwise to f and the set $\{f_n : n \geq 0\}$ is norm bounded i.e. $\{\|f_n\| : n \geq 0\} \subseteq \mathbb{R}$ is bounded.*

Proof. The $\Leftarrow$ direction is proven using the Dominated Convergence Theorem, while the $\Rightarrow$ direction is proven by choosing suitable Radon measures. □

Fact 3.35. *Suppose X is compact. Then weakly compact sets in $\mathrm{C}(X)$ are bounded.*

Proof. Let A be weakly compact in $\mathrm{C}(X)$. Towards a contradiction suppose A is not bounded. Then we can build an unbounded sequence $(f_n) \subseteq A$, with $\|f_n\| \leq \|f_{n+1}\|$. By Eberlein–Šmulian Theorem we can extract a subsequence (f_{n_k}) which converges weakly to some element in $\mathrm{C}(X)$. By Fact 3.34 this is not possible as $\{f_{n_k} : k \geq 0\}$ must be norm bounded. □

Remark 3.36. *The preceding fact is actually true in a more general context. The weakly compact subsets of any Banach space E are norm bounded. One standard way of proving this is using the Hahn-Banach Theorem and the Uniform Boundedness Principle. The Uniform Boundedness Principle has a simple proof from the Baire Category Theorem [Roy10, Sec. 13.5].*

Fact 3.37. *Suppose X is compact and let $A \subseteq \mathrm{C}(X)$ be weakly relatively compact. Then A is relatively compact in $\mathrm{C}_s(X)$.*

Proof. By the Eberlein–Šmulian Theorem, A is weakly relatively sequentially compact. Using Fact 3.34 we obtain that A is relatively sequentially compact in $\mathrm{C}_s(X)$ and by Theorem 3.30, A is relatively compact in $\mathrm{C}_s(X)$. □

Theorem 3.38. *[Gro52, Theorem 5] Suppose X is a compact space. For a subset A of $\mathrm{C}(X)$ to be weakly relatively compact, it is necessary and sufficient that it is bounded and relatively compact in $\mathrm{C}_s(X)$.*

Proof. The necessity is demonstrated by the facts mentioned earlier. For the sufficiency, suppose A is bounded and relatively compact in $\mathrm{C}_s(X)$. Then by Eberlein–Šmulian Theorem, it is enough to prove that from any sequence (f_n) of A we can extract a weakly convergent subsequence, which by Fact 3.34 is equivalent to finding a pointwise convergent subsequence of (f_n). Finally, by Theorem 3.30, (f_n) has a pointwise convergent subsequence. □

The next remark is used in the proof of Theorem 3.40 below.

Remark 3.39. *To any $f \in \mathrm{C}_b(X)$ we assign $I_f = \{x \in \mathbb{C} : |x| \leq \|f\|\}$. Then consider the map $e : X \to \prod_{f \in \mathrm{C}_b(X)} I_f$ defined by $x \mapsto (f(x))_{f \in \mathrm{C}_b(X)}$. The closure of $e(X)$ in $\prod_{f \in \mathrm{C}_b(X)} I_f$, denoted by βX, is called the Stone-Čech compactification of X (it can be defined in many more ways, for example using real-valued functions instead of complex-valued [Wil12, p.137]). It has the important property that for any continuous map $f : X \to K$, where K is compact and Hausdorff, exists a unique $F : \beta X \to K$, denoted by $\mathfrak{e}(f)$, such that $f = F \circ e$. Since βX is the closure of $e(X)$, $e(X)$ is dense in βX. Moreover, the map $e : X \to e(X)$ is an embedding if and only if X is a Tychonoff space [Wil12, p.136-137]).*

Therefore there exists a bijective correspondence $\mathfrak{e} : \mathrm{C}_b(X) \to \mathrm{C}(\beta X)$. For $f, g \in \mathrm{C}_b(X)$ we have that $\|\mathfrak{e}(f) - \mathfrak{e}(g)\| = \|f - g\|$. Hence $\mathfrak{e}$ is a homeomorphism.

Theorem 3.40. *[Gro52, Theorem 6] Let X_0 be a dense subset of X. A subset A of $\mathrm{C}_b(X)$ is weakly relatively compact if and only if:*

(i) A is bounded, and

(ii) for any two sequences $(f_n) \subseteq A$ and $(x_m) \subseteq X_0$ we have

$$\lim_m \lim_n f_n(x_m) = \lim_n \lim_m f_n(x_m)$$

provided both limits exist.

Proof. If X is Hausdorff and compact, then by Theorem 3.38, the relative weak compactness of A is equivalent to A being bounded and relatively compact in $\mathrm{C}_s(X)$. Furthermore, assuming A is bounded, by Corollary 3.28, the relative compactness of A in $\mathrm{C}_s(E)$ is equivalent condition (ii), thus proving the theorem in this case.

For X a general topological space, we use βX to translate the problem to the compact and Hausdorff setting. Since $\mathfrak{e}$ is a homeomorphism between $\mathrm{C}_b(X)$ and $\mathrm{C}(\beta X)$, we have that A weakly relatively compact in $\mathrm{C}_b(X)$ is equivalent to $\mathfrak{e}(A)$ weakly relatively compact in $\mathrm{C}(\beta X)$. Clearly, A bounded is equivalent to $\mathfrak{e}(A)$ bounded.

Since $e(X)$ is dense in βX, it follows that $e(X_0)$ is dense in βX. Thus, it remains to verify that condition (ii) is also equivalent to condition (ii) with X_0, A replaced by $e(X_0), \mathfrak{e}(A)$, respectively; this is easy to do using Remark 3.39. □

3.3 Topological Proof of the Fundamental Theorem of Stability

Throughout this subsection let φ be an $\mathcal{L}(x, y)$-formula, T be an $\mathcal{L}$-theory, and $\boldsymbol{M} \models T$. Before showing the topological proof of the fundamental theorem of stability we need another characterization of stable formulas.

Proposition 3.41. *The formula φ is stable in $\boldsymbol{M}$ if and only if for all $(a_n) \subseteq M_x, (b_m) \subseteq M_y$ we have*

$$\lim_m \lim_n \varphi(a_n, b_m)^{\boldsymbol{M}} = \lim_n \lim_m \varphi(a_n, b_m)^{\boldsymbol{M}}$$

provided both limits exist.

Proof. Suppose φ is stable. If for some $(a_n) \subseteq M_x, (b_m) \subseteq M_y$ we have that both limits $\lim_m \lim_n \varphi(a_n, b_m)^{\boldsymbol{M}}, \lim_n \lim_m \varphi(a_n, b_m)^{\boldsymbol{M}}$ exist and differ by a nonzero amount

$$|\lim_m \lim_n \varphi(a_n, b_m)^{\boldsymbol{M}} - \lim_n \lim_m \varphi(a_n, b_m)^{\boldsymbol{M}}| = \varepsilon > 0,$$

then by similar arguments as in Remark 3.27 we can find subsequences $(a_{i_m})_m$ and $(b_{j_m})_m$ such that for all $m < n$ we have

$$|\varphi(a_{i_m}, b_{j_n}) - \varphi(a_{i_n}, b_{j_m})| \geq \varepsilon/2,$$

which contradicts the stability of φ. Conversely, suppose φ is not stable, and take $\varepsilon > 0$ such that φ is not ε-stable. Again, using arguments as in Remark 3.27 starting from sequences $(a_n) \subseteq M_x, (b_m) \subseteq M_y$ such that $|\varphi(a_m, b_n) - \varphi(a_n, b_m)| \geq \varepsilon$, we can extract subsequences $(a_{i_m})_m$ and $(b_{j_m})_m$ such that both limits $\lim_m \lim_n \varphi(a_{i_n}, b_{j_m})^{\boldsymbol{M}}$, $\lim_n \lim_m \varphi(a_{i_n}, b_{j_m})^{\boldsymbol{M}}$ exist and differ by ε, contradiction. $\square$

Now, we can show the topological proof as in [BY14].

Proof of Theorem 3.7. Fix a type $p \in \mathrm{S}_\varphi(M)$ and let $X = \mathrm{S}_{\varphi^{op}}(M)$, $X_0 = \{\mathrm{tp}_{\varphi^{op}}(b|M) : b \in M_y\}$ be the set of realized types in $\mathrm{S}_{\varphi^{op}}(M)$, and $A = \{\varphi(\dot{a}, y) : a \in M_x\}$. Clearly X_0 is dense in X and $\mathrm{C}_b(X) = \mathrm{C}(X)$, as X is compact. By Theorem 3.40, φ is stable if and only if A is relatively weakly compact in $\mathrm{C}(X)$.

Now, let $(a_i)_{i \in I} \subseteq M_x$ be a net such that $\mathrm{tp}_\varphi(a_i|M)$ converges to p. Since A is relatively weakly compact and nets in compact spaces admit convergent subnets, we can assume $\varphi(\dot{a}_i, y)_{i \in I}$ converges weakly to some $\Phi \in \mathrm{C}(X)$. As weak convergence implies pointwise convergence, we obtain $\Phi \in \mathrm{C}(\mathrm{S}_{\varphi^{op}}(M), [0,1])$, therefore Φ is an M-definable φ^{op}-predicate. Finally, for every $b \in M_y$ we have

$$\varphi(x, b)^p = \lim_{i \in I} \varphi(a_i, b)^{\boldsymbol{M}} = \Phi(b),$$

showing that p is definable by the M-definable φ-predicate Φ. $\square$

And the fundamental theorem of stability Theorem 3.14, which is also Corollary 4 in [BY14], is proven similarly as in Theorem 3.14. Next, we present a final remark on the proofs of The Fundamental Theorem of Stability.

Remark 3.42. *Observe that this proof is more complicated, using non-trivial topological results, rather than the simple proof of Theorem 3.7 presented in the section " Basics of Stability Theory". The reason is that this topological proof, presented as in [BY14, Theorem 3], and the proof presented in [BYU10, Proposition 7.7] both use the strategy of approximating the function $d_\varphi p$. This approach seems to be harder as both of these proofs either use non-trivial topological facts, or Ramsey's Theorem. It turns out, as presented in the first section of this chapter, that it is much easier to just define $d_\varphi p$ and directly prove that it is continuous, a proof which only uses the definition of types and no additional theorems.*

Recently there appeared a new article [BYI25] which in (Theorem 2.7) gives a new proof of Theorem 3.7. Their proof is based on a functional analysis approach and uses the Downward Löwenheim-Skolem Theorem to reduce the proof to a separable elementary substructure of $\boldsymbol{M}$.

3.4 Additional Results

Inspired by the results mentioned in the articles [BY14] and [BYI25], in this section we state some more additional consequences of stability of a formula.

Throughout this section let φ be an $\mathcal{L}(x, y)$-formula and $p \in \mathrm{S}_\varphi(M)$.

Definition 3.43. Suppose φ is stable. In order to not create confusion between the type p and the function $p : \mathrm{S}_{\varphi^{op}}(M) \to [0, 1]$ defined in expression (3.1), we will denote this function $p : \mathrm{S}_{\varphi^{op}}(M) \to [0, 1]$ by $\dot{p}$. Hence, for example for $q \in \mathrm{S}_{\varphi^{op}}(M)$ we would have $\dot{p}(q) = \dot{q}(p)$, by expression 3.5. This notation is also natural, as seen by the discussion before Definition 3.8. Moreover it is clear that $\dot{p} = d_\varphi p$.

Suppose φ is stable. By Lemma 3.12 (namely inequality (3.4) and the way the function $\dot{p}$ is built, see expression (3.1)), results that for every $\delta > 0$ and $q \in \mathrm{S}_{\varphi^{op}}(M)$ there exists an open set $p \in O \subseteq \mathrm{S}_\varphi(M)$ such that for any $a \in M_x$ with $\mathrm{tp}_\varphi(a) \in O$ we have

$$|a^q - \dot{p}(q)| \leq \delta.$$

Therefore if we have a net $(\mathrm{tp}(a_i))_{i \in I} \to p$, then $(\varphi(\dot{a_i}, y)(q))_{i \in I}$ converges (in $[0, 1]$) to $p(q)$. Since, this is true for any q,

$$(\varphi(\dot{a_i}, y))_{i \in I} \text{ converges pointwise to } \dot{p}.$$

This is very interesting as $\dot{p}$ is a definable φ-predicate, hence it is a limit of φ-formulas, and also $\dot{p}$ is a pointwise limit, not necessarily a uniform limit, of the net $(\varphi(\dot{a_i}, y))_{i \in I}$. Even though $(\varphi(\dot{a_i}, y))_{i \in I}$ converges pointwise to $\dot{p}$, we can't use Fact 3.34 to state that $(\varphi(\dot{a_i}, y))_{i \in I}$ converges weakly to $\dot{p}$, as the the fact only applies when the net is a sequence.

Remark 3.44. *[BY14, Theorem 5] Suppose φ is stable in $\boldsymbol{M}$. Since we only need a countable number of parameters to define the definable predicate $d_\varphi p$ on $\boldsymbol{M}$, let $\boldsymbol{N} \preccurlyeq \boldsymbol{M}$ be a separable elementary substructure containing the parameters defining $d_\varphi p$ and containing the parameters from the $\mathcal{L}$-formula φ. It is clear that $d_\varphi(p|_N)$ is the definable predicate corresponding to $d_\varphi p$. Moreover, we observe that since $\boldsymbol{N}$ is separable, every point from $\mathrm{S}_\varphi(N)$ has a countable basis of neighborhoods (here we simply use a countable dense subset of $\boldsymbol{N}$). In particular there exists a sequence $(a_n) \subseteq N_x$ such that $\mathrm{tp}_\varphi(a_n|N) \to p|_N$ in the logic topology. But then, by the previous paragraph the sequence of functions $\varphi(\dot{a_n}, y) : \mathrm{S}_{\varphi^{op}}(N) \to [0, 1]$ converges pointwise to $d_\varphi(p|_N)$. Then, clearly, also the sequence of functions $(\varphi(\dot{a_n}, y)) : \mathrm{S}_{\varphi^{op}}(M) \to [0, 1]$, abusing notation, also converges pointwise to $d_\varphi p$. Summarizing, we have proved that $d_\varphi p$ is the weak limit of the sequence $(\varphi(\dot{a_n}, y))_n$.*

Additionally, for $b \in M_y$ note that $(\varphi(\dot{a_n}, y))_n$ converging weakly to $d_\varphi p$ implies that

$$\lim_n \varphi(a_n, b) = \varphi(x, b)^p.$$

Therefore, p is the limit of the countable sequence $(\mathrm{tp}_\varphi(a_n|M))_n$ in $\mathrm{S}_\varphi(M)$, with respect to the logic topology. In particular, p is the limit of a countable sequence of realized types.

By verifying the proof of Lemma 3.9, we have that:

Lemma 3.45. *Suppose φ is ε-stable in $\boldsymbol{M}$, and let p be a φ-type and q be a φ^{op}-type. Then there exist $a_1, \ldots, a_n \in M_x$ such that for any $b, b' \in M_y$ satisfying the $\mathcal{L}_M$-condition $\max(|\varphi(a_1, y) - a_1^q|, \ldots, |\varphi(a_n, y) - a_n^q|) \leq \varepsilon/2$ we have*

$$|b^p - (b')^p| \leq 4\varepsilon.$$

For every $q \in \mathrm{S}_{\varphi^{op}}(M)$ we pick $(a_i)_{i \in Fin_q} \subseteq M_x$ satisfying the conclusion of Lemma 3.45, where Fin_q is a finite set of elements. By compactness of $\mathrm{S}_{\varphi^{op}}(M)$ there exist $q_1, \ldots, q_n \in \mathrm{S}_{\varphi^{op}}(M)$ such that

$$\bigcup_{1 \leq m \leq n} [\max((\varphi(a_i, y) - a_i^{q_m})_{i \in Fin_{q_m}}) \leq \varepsilon/4] = \mathrm{S}_{\varphi^{op}}(M).$$

Let $Fin = \cup_m Fin_{q_m}$ and suppose $b, b' \in M_y$ are such that $|\varphi(a_i, b) - \varphi(a_i, b')| \leq \varepsilon/4$ for each $i \in Fin$. Then, since

$$\mathrm{tp}(b) \in [\max((\varphi(a_i, y) - a_i^{q_m})_{i \in Fin_{q_m}}) \leq \varepsilon/4],$$

for some m, $b' \models \max((\varphi(a_i, y) - a_i^{q_m})_{i \in Fin_{q_m}}) \leq \varepsilon/2$, and by Lemma 3.45,

$$|b^p - (b')^p| \leq 4\varepsilon.$$

Summarizing we have just proved that:

Proposition 3.46. *Suppose φ is ε-stable in $\boldsymbol{M}$. Then there exist elements $a_1, \ldots, a_n \in M_x$ such that the following holds:*

$$\text{for every } 1 \leq m \leq n \text{ if } |\varphi(a_m, b) - \varphi(a_m, b')| \leq \varepsilon/4, \text{ then } |b^p - (b')^p| \leq 4\varepsilon.$$

This proposition is a weaker variant, in terms of bounds, of Lemma B.2 from [BYU10]. They prove Lemma B.2 using the countably infinite Ramsey theorem, and it is the main result they use to uniformly approximate $d_\varphi p$ using φ-formulas, thereby proving $d_\varphi p$ is a φ-predicate. Inspired by Lemma 3.12 we define:

Definition 3.47. We call φ to be *approximately stable in* $\boldsymbol{M}$ if for any $p \in \mathrm{S}_\varphi(M)$, $q \in \mathrm{S}_{\varphi^{op}}(M)$ and for any $\delta > 0$ there exists $\varepsilon > 0$ and $a_1, \ldots, a_n \in M_x, b_1 \ldots, b_n \in M_y$ such that for any $a \in M_x$ and $b \in M_y$ satisfying

$$a \models \max(|\varphi(x, b_1) - b_1^p|, \ldots, |\varphi(x, b_n) - b_n^p|) \leq \varepsilon$$
$$b \models \max(|\varphi(a_1, y) - a_1^q|, \ldots, |\varphi(a_n, y) - a_n^q|) \leq \varepsilon$$

we have

$$|a^q - b^p| \leq \delta.$$

Definition 3.48. We define $\mathrm{R}_\varphi(M)$ to be the subset of $\mathrm{S}_\varphi(M)$ consisting of the realized φ-types in $\boldsymbol{M}$.

We note that for φ (not necessarily stable), we can build the function $\varphi : \mathrm{R}_\varphi(M) \times \mathrm{R}_{\varphi^{op}}(M) \to [0,1]$ defined by $\varphi(p,q) = \varphi^{\boldsymbol{M}}(a,b)$ where $a \in M_x$, $b \in M_y$ realize p, q, respectively.

Theorem 3.49. *The following are equivalent:*

(i) φ *is stable in* $\boldsymbol{M}$*;*

(ii) φ *is approximately stable in* $\boldsymbol{M}$*;*

(iii) every $p \in \mathrm{S}_\varphi(M)$ *is definable by a* φ^{op}*-predicate* $\dot{p}$*, every* $q \in \mathrm{S}_{\varphi^{op}}(M)$ *is definable by a* φ*-predicate* $\dot{q}$*, and moreover, the "symmetry property" holds* $\dot{p}(q) = \dot{q}(p)$*;*

(iv) The function $\varphi : \mathrm{R}_\varphi(M) \times \mathrm{R}_{\varphi^{op}}(M) \to [0,1]$ *can be extended to a function* $\varphi : \mathrm{S}_\varphi(M) \times \mathrm{S}_{\varphi^{op}}(M) \to [0,1]$ *such that for every* $p \in \mathrm{S}_\varphi(M)$ *and* $q \in \mathrm{S}_{\varphi^{op}}(M)$ *both of* $\varphi(p,\cdot) : \mathrm{S}_{\varphi^{op}}(M) \to [0,1]$ *and* $\varphi(\cdot,q) : \mathrm{S}_\varphi(M) \to [0,1]$ *are continuous.*

Proof. The implication $(i) \implies (ii)$ holds by Lemma 3.12, the implication $(ii) \implies (iii)$ is proved in the paragraph after Lemma 3.12, and the implication $(iii) \implies (iv)$ is obvious.

Suppose (iv). We will prove the double limit characterization for φ stated in Proposition 3.41, which in turn proves φ is stable in $\boldsymbol{M}$. Hence, let $(a_n) \subseteq M_x, (b_m) \subseteq M_y$ be two sequences for which $\lim_m \lim_n \varphi(a_n,b_m)^{\boldsymbol{M}}$ and $\lim_n \lim_m \varphi(a_n,b_m)^{\boldsymbol{M}}$ exist. Since $\mathrm{S}_\varphi(M)$ and $\mathrm{S}_{\varphi^{op}}(M)$ are compact topological spaces, we can take accumulation points $p \in \mathrm{S}_\varphi(M)$ and $q \in \mathrm{S}_{\varphi^{op}}(M)$ of the sequences $(\mathrm{tp}_\varphi(a_n|M))_n$ and $(\mathrm{tp}_{\varphi^{op}}(b_n|M))_n$, respectively. Then, since the single limits exist we note that

$$\lim_m \lim_n \varphi(a_n,b_m)^{\boldsymbol{M}} = \lim_m \varphi(p,b_m) = \varphi(p,q),$$

and similarly

$$\lim_n \lim_m \varphi(a_n,b_m)^{\boldsymbol{M}} = \lim_n \varphi(a_n,q) = \varphi(p,q),$$

therefore we obtain that $\lim_n \lim_m \varphi(a_n,b_m)^{\boldsymbol{M}} = \lim_m \lim_n \varphi(a_n,b_m)^{\boldsymbol{M}}$. □

The equivalences $(i) \Leftrightarrow (iii)$ and $(i) \Leftrightarrow (iv)$ are also established, with a different approach, in Lemma 2.5 and Theorem 2.7 of the article [BYI25].

Next, I present an alternative proof of $(i) \implies (iv)$, communicated to me by Itaï Ben Yaacov, which is very similar in style to Lemma 3.12. We start with a lemma.

Lemma 3.50. *Suppose* φ *is stable in* $\boldsymbol{M}$*. Then for any nets* $(a_\lambda)_{\lambda\in\Lambda}$ *and* $(b_\gamma)_{\gamma\in\Gamma}$ *in* M_x, M_y*, respectively, we have*

$$\lim_{\lambda\in\Lambda} \lim_{\gamma\in\Gamma} \varphi(a_\lambda,b_\gamma)^{\boldsymbol{M}} = \lim_{\gamma\in\Gamma} \lim_{\lambda\in\Lambda} \varphi(a_\lambda,b_\gamma)^{\boldsymbol{M}},$$

provided both limits exist.

Proof. By Proposition 3.41 we know φ satisfies the double limit property for sequences. Suppose the contradiction, let $(a_\lambda)_{\lambda\in\Lambda}$ and $(b_\gamma)_{\gamma\in\Gamma}$ be two nets in M_x, M_y, such that both limits $\lim_{\lambda\in\Lambda}\lim_{\gamma\in\Gamma}\varphi(a_\lambda,b_\gamma)^{\boldsymbol{M}}, \lim_{\gamma\in\Gamma}\lim_{\lambda\in\Lambda}\varphi(a_\lambda,b_\gamma)^{\boldsymbol{M}}$ exist, and suppose

$$\lim_{\lambda\in\Lambda}\lim_{\gamma\in\Gamma}\varphi(a_\lambda,b_\gamma)^{\boldsymbol{M}} > r > s > \lim_{\gamma\in\Gamma}\lim_{\lambda\in\Lambda}\varphi(a_\lambda,b_\gamma)^{\boldsymbol{M}}, \tag{3.6}$$

the case $r < s$ is treated similarly. Since $\mathrm{S}_\varphi(M)$ and $\mathrm{S}_{\varphi^{op}}(M)$ are compact topological spaces, there obtain subnets Λ' of Λ and Γ' of Γ, with the corresponding maps $f:\Lambda'\to\Lambda, g:\Gamma'\to\Gamma$ which are order-preserving and cofinal, and types $p\in\mathrm{S}_\varphi(M), q\in\mathrm{S}_{\varphi^{op}}(M)$ such that $(\mathrm{tp}_\varphi(a_{f(\lambda')}|M))_{\lambda'\in\Lambda'}\to p$ and $(\mathrm{tp}_{\varphi^{op}}(b_{g(\gamma')}|M))_{\gamma'\in\Gamma'}\to q$.

Note that the expression in (3.6) does not change if instead of considering the nets $(a_\lambda)_{\lambda\in\Lambda}, (b_\gamma)_{\gamma\in\Gamma}$ we consider the nets $(a_{f(\lambda')})_{\lambda'\in\Lambda'}, (b_{g(\gamma')})_{\gamma'\in\Gamma'}$. Therefore we can suppose $(\mathrm{tp}_\varphi(a_\lambda|M))_{\lambda\in\Lambda}\to p$ and $(\mathrm{tp}_{\varphi^{op}}(b_\gamma|M))_{\gamma\in\Gamma}\to q$. Then, we obtain

$$\lim_{\lambda\in\Lambda}\varphi(a_\lambda,y)^q = \lim_{\lambda\in\Lambda}\lim_{\gamma\in\Gamma}\varphi(a_\lambda,b_\gamma)^{\boldsymbol{M}} > r > s > \lim_{\gamma\in\Gamma}\lim_{\lambda\in\Lambda}\varphi(a_\lambda,b_\gamma)^{\boldsymbol{M}} = \lim_{\gamma\in\Gamma}\varphi(x,b_\gamma)^p.$$

We now recursively construct two sequences $(a_n)\subseteq M_x$ and $(b_n)\subseteq M_y$ such that for each $n\in\mathbb{N}$, there exist $\lambda\in\Lambda$ and $\gamma\in\Gamma$ with $a_n = a_\lambda$ and $b_n = b_\gamma$. To begin, choose $a_0\in M_x$ and $b_0\in M_y$ such that $\varphi(a_0,y)^q > r$ and $s > \varphi(x,b_0)^p$. Suppose by induction that we have already selected $a_0,\dots,a_{n-1}$ and $b_0,\dots,b_{n-1}$ satisfying $\varphi(a_m,y)^q > r$ and $s > \varphi(x,b_m)^p$ for all $m\le n-1$. Then we can pick a_n and b_n such that $\varphi(a_n,y)^q > r$, $s > \varphi(x,b_n)^p$ and $\varphi(a_m,b_n) > r$ and $s > \varphi(a_n,b_m)$ for each $m\le n-1$. But then we obtain

$$\lim_n\lim_m\varphi(a_n,b_m)\ge r > s\ge\lim_m\lim_n\varphi(a_n,b_m),$$

contradicting the stability of φ. □

Proposition 3.51. *Suppose the formula φ is stable in $\boldsymbol{M}$. Then the map $\varphi:\mathrm{R}_\varphi(M)\times\mathrm{R}_{\varphi^{op}}(M)\to[0,1]$ can be extended to a map $\varphi:\mathrm{S}_\varphi(M)\times\mathrm{S}_{\varphi^{op}}(M)\to[0,1]$ such that for every $p\in\mathrm{S}_\varphi(M)$ and $q\in\mathrm{S}_{\varphi^{op}}(M)$ we have $\varphi(p,\cdot):\mathrm{S}_{\varphi^{op}}(M)\to[0,1]$ and $\varphi(\cdot,q):\mathrm{S}_\varphi(M)\to[0,1]$ are continuous. Moreover, this extension is unique.*

Proof. Let $p\in\mathrm{S}_\varphi(M)$ and $q\in\mathrm{S}_{\varphi^{op}}(M)$. Consider two nets $(a_\lambda)_{\lambda\in\Lambda}$ in M_x and $(b_\gamma)_{\gamma\in\Gamma}$ in M_y such that

$$\mathrm{tp}_\varphi(a_\lambda|M)\to p \text{ and } \mathrm{tp}_{\varphi^{op}}(b_\gamma|M)\to q.$$

We first observe that if both limits $\lim_\lambda\varphi(a_\lambda,y)^q, \lim_\gamma\varphi(x,b_\gamma)^p$ exist, then by stability of φ we have

$$\lim_\lambda\varphi(a_\lambda,y)^q = \lim_\lambda\lim_\gamma\varphi(a_\lambda,b_\gamma) = \lim_\gamma\lim_\lambda\varphi(a_\lambda,b_\gamma) = \lim_\gamma\varphi(x,b_\gamma)^p.$$

Suppose that the limit $\lim_\lambda\varphi(a_\lambda,y)^q$ does not exist. Then there exist two subnets Λ',Λ'' of Λ such that the limits $\lim_{\lambda'\in\Lambda'}\varphi(a_{\lambda'},y)^q, \lim_{\lambda''\in\Lambda''}\varphi(a_{\lambda''},y)^q$ exist and are not equal. We can also pick a subnet Γ' of Γ such that $\lim_{\gamma'\in\Gamma'}\varphi(x,b_{\gamma'})^p$ exists. But, then, by the argument in the previous paragraph

$$\lim_{\lambda'\in\Lambda'}\varphi(a_{\lambda'},y)^q = \lim_{\gamma'\in\Gamma'}\varphi(x,b_{\gamma'})^p = \lim_{\lambda''\in\Lambda''}\varphi(a_{\lambda''},y)^q,$$

a contradiction. Therefore, $\lim_\lambda \varphi(a_\lambda, y)^q$ exists. A similar argument shows that the value of $\lim_\lambda \varphi(a_\lambda, y)^q$ is independent of the choice of the net (a_λ) converging to p, and likewise the limit $\lim_\gamma \varphi(x, b_\gamma)^p$ exists and is independent of the choice of the net (b_γ) converging to q. Moreover both these limits are equal. Therefore we can define

$$\varphi(p, q) = \lim_\lambda \varphi(a_\lambda, y)^q = \lim_\gamma \varphi(x, b_\gamma)^p.$$

This way, for every p, q, we define $\varphi(p, q)$ and the map $\varphi : \mathrm{S}_\varphi(M) \times \mathrm{S}_{\varphi^{op}}(M) \to [0, 1]$ which is readily seen to extend $\varphi : \mathrm{R}_\varphi(M) \times \mathrm{R}_{\varphi^{op}}(M) \to [0, 1]$. To prove continuity, suppose $\varphi(\cdot, q)$ is not continuous at p. Then there exists a net $(p_\delta)_{\delta\in\Delta}$ (with Δ the directed set of neighborhoods of p ordered by reverse inclusion) such that $p_\delta \to p$ and $|\varphi(p_\delta, q) - \varphi(p, q)| > \varepsilon > 0$, for some fixed $\varepsilon > 0$. Then for every $\delta \in \Delta$, by using a net of realized types converging to p_δ and the way we extended φ, we can pick $c_\delta \in M_x$ such that $\mathrm{tp}_\varphi(c_\delta|M) \in \delta$ and $|\varphi(c_\delta, y)^q - \varphi(p, q)| > \varepsilon$. But, then since $\mathrm{tp}(c_\delta|M) \to p$ we must have $\lim_\delta \varphi(c_\delta, y)^q = \varphi(p, q)$, a contradiction. Similarly one shows that $\varphi(p, \cdot)$ is continuous. The moreover part of the proposition follows from the fact that the realized φ-types and the realized φ^{op}-types are dense in their respective φ-type spaces. $\square$

4 Applications to Probability

This chapter uses extensively the article [BH23] in order to present the results on the continuous model theory of Probability Algebras. The exposition develops progressively towards establishing that the theory of Atomless Probability Algebras is $\aleph_0$-stable. The reader is assumed to have basic knowledge of boolean algebras, probability theory and topology.

4.1 Probability Algebras

A *probability algebra* is a tuple $\boldsymbol{B} = (B, 0, 1, \cdot^c, \wedge, \vee, \mathbb{P})$ with $(B, 0, 1, \cdot^c, \wedge, \vee)$ a σ-complete boolean algebra (every countable set of elements has a supremum and infimum) and $\mathbb{P}$ is a strictly positive probability measure on B (for every $a \in B, 0 < a \implies 0 < \mathbb{P}(a)$). Every probability algebra comes equipped with a natural metric $d^{\boldsymbol{B}}(a, b) = d(a, b) = \mathbb{P}(a \triangle b)$ for every $a, b \in B$. The next result uses calculations similar to those in [Fre11, Lemma 323F].

Theorem 4.1. *Let $\boldsymbol{B} = (B, 0, 1, \cdot^c, \wedge, \vee, \mathbb{P})$ be a boolean algebra where $\mathbb{P}$ is a finitely additive probability measure. Then $\boldsymbol{B}$ is a probability algebra if and only if $(B, d^{\boldsymbol{B}})$ is a complete metric space.*

Proof. Suppose $\boldsymbol{B}$ is a probability algebra. Let (a_n) be a cauchy sequence in B such that $d^{\boldsymbol{B}}(a_n, a_{n+1}) \leq \frac{1}{2^n}$ for every n. Let $a_{\text{inf}} = \vee_n (\wedge_{m \geq n} a_m)$ and $a_{\text{sup}} = \wedge_n (\vee_{m \geq n} a_m)$. Then for every n, we have the relations

$$a_{\text{inf}} \triangle a_{\text{sup}} \leq (a_{\text{inf}} \triangle a_n) \cup (a_n \triangle a_{\text{sup}}) \leq \bigvee_{m \geq n} a_m \triangle a_{m+1} \quad \text{and} \quad \mathbb{P}\Big(\bigvee_{m \geq n} a_m \triangle a_{m+1} \Big) \leq \frac{1}{2^{n-1}}.$$

Hence, as $n \to 0$, we obtain $d^{\boldsymbol{B}}(a_{\text{inf}} \triangle a_{\text{sup}}) = 0$ and $\lim_n d^{\boldsymbol{B}}(a_{\text{inf}} \triangle a_n) = 0$. This implies that $a_{\text{inf}} = a_{\text{sup}}$, and this element is the cauchy limit of (a_n).

Conversely, suppose $(B, d^{\boldsymbol{B}})$ is a complete metric space. Since $d^{\boldsymbol{B}}$ is a metric, it follows that $\mathbb{P}$ is strictly positive. Let (a_n) be an increasing sequence of elements from B. Since $\lim \mathbb{P}(a_n)$ exists, (a_n) is a cauchy sequence in $d^{\boldsymbol{B}}$; let a be its limit. For any n, the element $a_n \vee a$ is also a limit of (a_n), which implies $a_n \vee a = a$, and hence $a_n \leq a$. Moreover, if $b \in B$ is an upper bound of (a_n), then $a \wedge b$ is a limit of (a_n), yielding $a = b$. This implies that B a σ-complete, and since $\lim \mathbb{P}(a_n) = \mathbb{P}(a)$, it follows that $\mathbb{P}$ is a strictly positive probability measure. $\square$

The standard method for obtaining probability algebras is by taking the quotient of a probability space with respect to the ideal of null sets. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability

space. Let $I \subseteq \mathcal{F}$ be the ideal of null sets with respect to the σ-algebra $\mathcal{F}$. We define the quotient σ-complete boolean algebra $\widehat{\mathcal{F}} = \mathcal{F}/I$ and denote by $\widehat{\mathbb{P}} : \widehat{\mathcal{F}} \to [0,1]$ the corresponding strictly positive probability measure. For an element $A \in \mathcal{F}$ we denote by $\widehat{A}$ its equivalence class in $\mathcal{F}\backslash I$. This induces the natural projection map $\widehat{(\cdot)} : \mathcal{F} \to \mathcal{F}\backslash I$. Then with the natural interpretations the tuple $\boldsymbol{B} = (\widehat{\mathcal{F}}, 0, 1, \cdot^c, \vee, \wedge, \widehat{\mathbb{P}})$ is a probability algebra, also called *the probability algebra of the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$.
We also have following trivial result:

Proposition 4.2. *[BH23, Lemma 2.2] Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space and* $\boldsymbol{B} = (\widehat{\mathcal{F}}, 0, 1, \vee, \wedge, \widehat{\mathbb{P}})$ *be its corresponding probability algebra. Then*

(i) For any σ*-subalgebra* $\mathcal{F}_0 \subseteq \mathcal{F}$*, in a natural way* $\boldsymbol{B}_0 = (\widehat{\mathcal{F}_0}, 0, 1, \cdot^c, \vee, \wedge, \widehat{\mathbb{P}})$ *is a probability subalgebra of* $\boldsymbol{B}$*;*

(ii) For any probability subalgebra $\boldsymbol{B}_0 \subseteq \boldsymbol{B}$ *there is a unique* σ*-subalgebra* $\mathcal{F}_0 \subseteq \mathcal{F}$ *which contains all the null sets of* $\mathcal{F}$ *and* $\boldsymbol{B}_0 = (\widehat{\mathcal{F}_0}, 0, 1, \cdot^c, \vee, \wedge, \widehat{\mathbb{P}})$*.*

We now present an important representation theorem for probability algebras.

Theorem 4.3. *[Fre11, Theorem 321J] Every probability algebra is isomorphic to a probability algebra of a probability space.*

In Theorem 4.5 of [BH23] they provide a model-theoretic proof of this using the Loeb measure construction. Here we will provide a proof similar in style to the proof of Theorem 321J of [Fre11].

Let $(B, 0, 1, \cdot^c, \vee, \wedge, \mu)$ be a boolean algebra equipped with a finitely-additive measure μ. By Stone's representation theorem for boolean algebras there is a boolean algebra isomorphism

$$[\cdot] : B \to \mathrm{Clopen}(\mathrm{St}(B)), a \to [a] = \{p : p \text{ an ultrafilter on } B, a \in p\}$$

where $\mathrm{St}(B)$ is the stone space of B, and $\mathrm{Clopen}(\cdot)$ is the boolean algebra of clopens of a topology. Hence we obtain a finitely additive measure on $\mathrm{Clopen}(\mathrm{St}(B))$, which by abuse of notation we denote by μ as well.

Definition 4.4. Let X be a topological space. We denote by $\mathrm{Borel}(X)$ the collection of all Borel sets of X, which is the smallest σ-algebra generated by the open sets. A *Radon measure* on X is a measure on $\mathrm{Borel}(X)$ which has *inner and outer regularity*, that is, for every Borel set A we have

$$\mu(A) = \sup\{\mu(K) : K \subseteq A, K \text{ is compact}\} = \inf\{\mu(U) : A \subseteq U, U \text{ is open}\},$$

and which is locally finite, that is, for every compact $K \subseteq X$, $\mu(K) < \infty$.

Our goal is to extend the finitely additive measure μ defined on $\mathrm{Clopen}(\mathrm{St}(B))$ to a measure, in particular a Radon measure, on the Borel space of $\mathrm{St}(B)$. From the theory of boolean algebras, we know that $\mathrm{St}(B)$ is a Hausdorff, compact and zero-dimensional space. By definition, a space is zero-dimensional if its set of clopens is a basis.

Proposition 4.5. *Let X be a hausdorff, compact, and zero-dimensional space with μ a finitely additive measure on* Clopen(X)*. Then μ extends to a Radon measure on X.*

Proof of Proposition 4.5. One way of proving this proposition is to first extend μ to the Baire space of X. The Baire space of X is the smallest σ-algebra on X for which all the continuous functions $X \to \mathbb{R}$ are measurable. This extension can be done using Carathéodory's extension theorem, because μ is a σ-measure on Clopen(X) (an easy consequence of compactness) and the Baire Space on X coincides with the σ-algebra generated by the clopens (a consequence of the fact that every continuous function $X \to [0,1]$ is a pointwise limit of continuous functions with finite image). The next step would be to apply [Dud18, Theorem 7.3.1], which states that any finite measure on the Baire space of a compact and hausdorff topological space can be extended to a measure on the corresponding Borel space. In the following steps, we present a proof similar to Theorem 7.3.1 from [Dud18], adapted for compact, Hausdorff, zero-dimensional spaces.

Define for any compact $K \subseteq X$,

$$\mu(K) = \inf\{\mu(C) \mid C \supseteq K \text{ clopen}\}.$$

Note that this agrees with the given finitely additive measure μ on Clopen(X).

1) Let K, L be two compact disjoint sets. Then $\mu(K \cup L) = \mu(K) + \mu(L)$. μ is monotone. And if $F_n \downarrow F$ is a decreasing sequence of compact sets then $\mu(F_n) \downarrow \mu(F)$.
Proof. The $\leq$ direction is obvious. Now for $\geq$, observe that there exist disjoint clopens A, B which separate K and L. Take any C clopen such that $K \cup L \subseteq C$. Then $K \subseteq C \cap A$ and $L \subseteq C \cap B$. And with the observation that, since μ is finitely additive on clopens,

$$\mu(C) \geq \mu(C \cap A) + \mu(C \cap B),$$

we can conclude the $\geq$ direction.
For the last part, just observe that for any clopen C such that $F \subseteq C$, by compactness there exists n such that $F_n \subseteq C$. □

2) For any $B \subseteq X$ define $\nu(B) = \inf\{\mu(K) \mid K \subseteq B \text{compact}\}$. If B is compact, then $\nu(B) = \mu(B)$, hence ν is an extension on $\mathcal{P}(X)$ of μ defined on all the compact sets. Additionally, we observe that ν is monotone.

Define $B \in \mathcal{M}$ iff for all $A \subseteq X$, $\nu(A) = \nu(A \cap B) + \nu(A \cap B^c)$. Call elements of $\mathcal{M}$ measurable. Obviously $\nu(A) \geq \nu(A \cap B) + \nu(A \cap B^c)$. So B is measurable if and only if $\nu(A) \leq \nu(A \cap B) + \nu(A \cap B^c)$ for all $A \subseteq X$. Since ν monotone and as a consequence of its definition, we can arrange A to be compact.

3) $\mathcal{M}$ is an algebra, ν is finitely additive on $\mathcal{M}$, and all compact sets belong to $\mathcal{M}$.
Proof. Clearly, $\mathcal{M}$ is closed under the complement operation. It is also closed under intersection, because for $B, C \in \mathcal{M}$ and every $A \subseteq X$ we have

$$\begin{aligned}\nu(A) = \nu(A \cap B) + \nu(A \cap B^c) &= \nu(A \cap B \cap C) + \nu(A \cap B \cap C^c) + \nu(A \cap B^c)\\ &= \nu(A \cap B \cap C) + \nu(A \cap (B \cap C)^c),\end{aligned}$$

where for the last equation we used $(A\cap(B\cap C)^c)\cap B = A\cap B\cap C^c$ and $(A\cap(B\cap C)^c)\cap B^c = A\cap B^c$. For finite additivity, if A, B are disjoint and measurable then

$$\nu(A\cup B) = \nu((A\cup B)\cap B) + \nu((A\cup B)\cap B^c) = \nu(B) + \nu(A).$$

Let K be a compact set. We have to prove that for any compact set L

$$\nu(L) \leq \nu(L\cap K) + \nu(L\cap K^c).$$

Take a clopen C such that $L\cap K \subseteq C$ and $\nu(C) - \nu(L\cap K) < \varepsilon$; then

$$\nu(L) \underset{\text{by Step 1}}{=} \nu(L\cap C) + \nu(L\backslash C) \leq \nu(C) + \nu(L\backslash K) \leq \nu(L\cap K) + \varepsilon + \nu(L\backslash K),$$

letting ε go to zero, we obtain the required result. □

4) Let $B_n \downarrow B$, be a decreasing sequence of subsets of X. Then $\nu(B_n) \downarrow \nu(B)$.
Proof. Suppose not, then fix $\nu(B) < \alpha \leq \nu(B_n)$ for each n. Let $\varepsilon > 0$ such that $0 < 2\varepsilon < \alpha - \nu(B)$. Then, by definition, for each n there exists a compact $K_n \subseteq B_n$ such that $\alpha - \varepsilon \leq \nu(K_n)$ and $\nu(B_n\backslash K_n) \leq \frac{\varepsilon}{2^n}$. Let $F_n = \bigcap_{i=1}^n K_i$. Then by the finite additivity of μ on $\mathcal{M}$ we get

$$\nu(K_n) \leq \nu(K_1\cap K_2\cap\cdots\cap K_n) + \sum_{i\in\overline{1,n}} \nu(K_n\backslash K_i) \leq \nu(K_1\cap K_2\cap\cdots\cap K_n) + \varepsilon$$

therefore

$$0 < \alpha - 2\varepsilon \leq \nu(K_1\cap K_2\cap\cdots\cap K_n) = \nu(F_n)$$

so $F_n \neq \emptyset$ and is compact. Let $F = \bigcap_n F_n$. Then $F_n \downarrow F$, $F \subseteq B$, F is compact, and hence by Step 1 we obtain that $\nu(F_n) \downarrow \nu(F)$. So $\alpha - 2\varepsilon \leq \nu(F) \leq \nu(B) < \alpha - 2\varepsilon$, a contradiction. □

5) $\mathcal{M}$ is a σ-algebra and ν is a measure on it.
Proof. Let E_n be increasing measurable sets with $E_n \uparrow E$. Then $\nu(E) = \nu(E_n) + \nu(E\cap E_n^c)$. Since $E\cap E_n^c \downarrow \emptyset$, by Step 4 we obtain $\nu(E\cap E_n^c) \downarrow 0$, so $\nu(E_n) \uparrow \nu(E)$. Fix A any set. Then

$$\nu(A) = \nu(A\cap E_n) + \nu(A\cap E_n^c) \leq \nu(A\cap E) + \nu(A\cap E_n^c),$$

and passing to the limit $n\to\infty$, using Step 4, we get

$$\nu(A) = \nu(A\cap E) + \nu(A\cap E^c),$$

obtaining E is measurable. Since $\mathcal{M}$ is an algebra and closed under unions of increasing sequences, it follows $\mathcal{M}$ is a σ-algebra. Moreover, because $\mathcal{M}$ contains all compacts, hence it also contains the σ-algebra $\mathrm{Borel}(X)$. □

6) Finally, it is obvious by construction, that ν is a Radon measure and that it is the unique Radon extension of μ to $\mathrm{Borel}(X)$. This yields a bijective correspondence between finitely additive measures on the boolean algebra of clopen subsets of X and the Radon measures on $\mathrm{Borel}(X)$. □

Proof of Theorem 4.3. Let $\boldsymbol{B} = (B, 0, 1, \cdot^c, , \vee, \wedge, \mathbb{P})$ be a probability algebra. Let $[\cdot] : B \to \mathrm{Clopen}(\mathrm{St}(B))$ be the natural isomorphism of boolean algebras and let μ be the corresponding finitely additive measure on $\mathrm{Clopen}(\mathrm{St}(B))$. While it is true that $[\cdot]$ is a measure-preserving boolean algebra isomorphism

$$\boldsymbol{B} \to (\mathrm{Clopen}(\mathrm{St}(B)), \emptyset, \mathrm{St}(B), \cdot^c, \cup, \cap, \mu),$$

it is not necessary that it is a σ-complete boolean algebra isomorphism. By Theorem 4.5 let ν be the Borel measure extending μ. Consider an increasing sequence (a_n) of elements of B with $a = \vee_n a_n$. Then since $\lim_n \mathbb{P}(a_n) = \mathbb{P}(a)$, we have $\lim_n \mu([a_n]) = \mu[a]$. This implies that $\nu([a] \setminus \cup_n [a_n]) = 0$. Based on this observation we see that the algebra $\mathcal{C}$ generated by $\mathrm{Clopen}(\mathrm{St}(B)) \cup \mathcal{N}_\nu$, with $\mathcal{N}_\nu$ the ideal of null sets of the measure space $(\mathrm{St}(B), \mathrm{Borel}(\mathrm{St(B)}), \nu)$, is a σ-algebra, and that $(\widehat{\mathcal{C}}, 0, 1, \cdot^c, \vee, \wedge, \widehat{\nu})$ is a probability algebra isomorphic to $\boldsymbol{B}$ via the isomorphism $\widehat{(\cdot)} \circ [\cdot]$. □

4.2 Model Theory of Probability Algebras

Let $\mathcal{L} = \mathcal{L}_{Pr} = \{0, 1, \cdot^c, \wedge, \vee, \mathbb{P}\}$ be a language, where $0, 1$ are constant symbols, $\cdot^c$ is a unary function symbol, $\wedge, \vee$ are binary functions symbols, and $\mathbb{P}$ is a unary relation symbol. The moduli of continuity are given by $\Delta_{\cdot^c}(\varepsilon) = \varepsilon, \Delta_\wedge(\varepsilon) = \varepsilon/2, \Delta_\vee(\varepsilon) = \varepsilon/2$ and $\Delta_{\mathbb{P}}(\varepsilon) = \varepsilon$ for every $\varepsilon > 0$ (according to Definition 1.6). Let Pr be the theory given by the following axioms [BH23, p.11]:

(i) Boolean algebra axioms: every boolean algebra axiom [Hal09, p.10] is an equality between terms and can therefore be expressed as an axiom in continuous logic. For example $\forall x(x \vee x = x)$ corresponds to $\sup_x d(x \vee x, x) = 0$ (see item (iii)).

(ii) Measure axioms:
$\mathbb{P}(0) = 0$ and $\mathbb{P}(1) = 1$,
$\sup_x \sup_y (\mathbb{P}(x \cap y) \dot{-} \mathbb{P}(x)) = 0$,
$\sup_x \sup_y (\mathbb{P}(x) \dot{-} \mathbb{P}(x \cup y)) = 0$,
$\sup_x \sup_y |(\mathbb{P}(x) \dot{-} \mathbb{P}(x \cap y)) - (\mathbb{P}(x \cup y) \dot{-} \mathbb{P}(y))| = 0$.

(iii) The axiom for the distance function:

$$\sup_x \sup_y |d(x, y) - \mathbb{P}(x \triangle y)| = 0.$$

Clearly, any probability algebra satisfies Pr. Conversely, let $\boldsymbol{M} \models Pr$. Then $\boldsymbol{M}$ is a boolean algebra with $\mathbb{P}$ a finitely-additive probability measure, and $(M, d^{\boldsymbol{M}})$ is a complete metric space. By Theorem 4.1 it follows that $\boldsymbol{M}$ is a probability algebra. Applying Theorem 4.3, we thus obtain:

Theorem 4.6. *[BH23, Theorems 4.3,4.5] Let $\boldsymbol{M}$ be an $\mathcal{L}_{Pr}$-structure. Then the following are equivalent:*

(i) $\boldsymbol{M}$ *is a model of* Pr*;*

(ii) $\boldsymbol{M}$ *is a probability algebra;*

(iii) $\boldsymbol{M}$ *is isomorphic to a probability algebra of a probability space.*

Throughout the end of this section we let $\boldsymbol{M}$ be a probability algebra. Before stating more structural results about the theory Pr, we first review some further fundamental concepts and definitions related to probability algebras.

Remark 4.7. *For precision, interpretations of symbols and formulas from the language* $\mathcal{L}_{Pr}$ *should be denoted by* $(\cdot)^{\boldsymbol{M}}$*. However, to maintain readability and clarity, we will generally write symbols as they appear, using* $(\cdot)^{\boldsymbol{M}}$ *only when necessary—particularly in contexts involving multiple models of* Pr *where ambiguity might arise.*

Definition 4.8. Let $a \in M$. We call a an *atom* if $a \neq 0$ and there is no $b < a$ with $\mathbb{P}(b) > 0$. We call a *atomless* if $a \neq 0$ and there is no atom $\leq a$. We call a *atomic* if $a \neq 0$ and there is no atomless element $\leq a$. The probability algebra $\boldsymbol{M}$ is called *atomless* if it has no atoms. We call a finite set of disjoint elements $P = \{a_1, \dots, a_n\} \subseteq M$ a *partition of* a if $a = \vee_m a_m$. The *diameter* of P is defined as $|P| = \max(\mathbb{P}(a_1), \dots, \mathbb{P}(a_n))$, and we say $\mathbb{P}$ is a ε*-partition* if $|P| \leq \varepsilon$. Given partitions $P_1, \dots, P_n$ of a we define their intersection to be the partition

$$\{a_1 \wedge \cdots \wedge a_n : a_1 \in P_1, \dots, a_n \in P_n\},$$

which is also a partition of a.

Note that the number of atoms in $\boldsymbol{M}$ is at most countable. Let A denote the set of atoms of $\boldsymbol{M}$. Since each atom has positive measure and they are pairwise disjoint, the set A can be enumerated as $a_1, a_2, \dots$ in such a way that

$$\mathbb{P}(a_1) \geq \mathbb{P}(a_2) \geq \dots .$$

Definition 4.9. The element $1_{\mathrm{at}}^{\boldsymbol{M}} = 1_{\mathrm{at}} = \bigvee_{a \in A} a$ is called the *atomic part* of M (if there are no atoms then we set $1_{\mathrm{at}} = 0$), while $1_{\mathrm{al}}^{\boldsymbol{M}} = 1_{\mathrm{al}} = (1_{\mathrm{at}})^c$ is called the *atomless part* of M.

More generally, for any $a \in M$, we define the *atomic part* of a, denoted a_{at}, to be the union of all atoms less than or equal to a. The *atomless part* of a, denoted a_{al}, is defined by $a_{\mathrm{al}} = a \setminus a_{\mathrm{at}}$.

Definition 4.10. Let $a \in M$. A representation of the form $a_{at} = \bigvee a_n$ with a_n an atom or zero, such that $(\mathbb{P}(a_n))_n$ is a decreasing sequence, is called an *atomic representation* of a. Clearly a can have more than one atomic representation, if some atoms have the same value.

It is clear that 1_{al} is atomless, and that every atomless element of M is $\leq 1_{\mathrm{al}}$. Atomless elements have the following interesting property.

Lemma 4.11. *Let $b \in M$ be atomless. Then for every $\varepsilon > 0$ there is an ε-partition of b.*

Proof. We may assume $b = 1$. Suppose the contrary. Since every ε-partition of 1 is also an ε'-partition of 1 for each $\varepsilon' \geq \varepsilon$, there exists $\varepsilon \in (0,1)$ such that there is no ε-partition of 1, but there exists an ε'-partition of 1 for each $\varepsilon' > \varepsilon$.

Let $P = \{a_1, \ldots, a_n\}$ be a $3\varepsilon/2$-partition of 1, with the minimal amount of elements of measure $> \varepsilon$. Suppose $\mathbb{P}(a_1) > \varepsilon$ and let $I = \{\mathbb{P}(a) : a \leq a_1, \mathbb{P}(a) \geq \varepsilon\}$. Let $\varepsilon' = \inf I$. Then for every n exists $b_n \leq a_1$ such that $\mathbb{P}(b_n) \leq \varepsilon' + 1/n$. Observe that by construction of P, any partition of a_1 must have an element of measure $> \varepsilon$. Therefore by taking the intersection of the partitions $\{b_1, a_1 \backslash b_1\}, \ldots, \{b_n, a_1 \backslash b_n\}$, we obtain that $c_n = b_1 \wedge \cdots \wedge b_n$ has measure $> \varepsilon$ and obviously $\mathbb{P}(c_n) \leq \varepsilon' + 1/n$. Then, for $c = \wedge_n c_n$ we have $\mathbb{P}(c) = \varepsilon'$. Since c is atomless, by taking any partition of $c = a \vee b$, it follows that $\mathbb{P}(a) < \varepsilon$ and $\mathbb{P}(b) < \varepsilon$, obtaining $a_1 = a \vee b \vee (a_1 \backslash c)$ which is an ε-partition of a_1, contradiction. $\square$

As an easy consequence, we have the following corollary.

Corollary 4.12. *Let $b \in M$ be atomless. Then $\{\mathbb{P}(a) : a \leq b\} = [0, \mathbb{P}(b)]$.*

Now, we will introduce some $\mathcal{L}_{Pr}$-formulas which will help us identifying whether an element is an atom or atomless, and find some special definable sets. We also introduce

$$A_1^{\boldsymbol{M}} = A_1 = \{a \in M : a = 0 \text{ or } a \text{ is an atom}\}.$$

Note that $1_{\mathrm{at}} = \bigvee_{a \in A_1} a$.

Definition 4.13. [BH23, p. 13] We define the following $\mathcal{L}_{Pr}$-formulas:

$$\chi(x) = \inf_y |\mathbb{P}(x \wedge y) - \mathbb{P}(x \wedge y^c)|, \qquad \varphi(z) = \mathbb{P}(x) \dot{-} \chi(x),$$
$$\varphi_1(x) = \inf_z (d(x,z) + \varphi(z)), \qquad \theta(x) = \sup_y \inf_z |\mathbb{P}(x \cap y \cap z) - \mathbb{P}(x \cap y \cap z^c)|.$$

Proposition 4.14. *[BH23, Proposition 4.7] Let $a \in M$. Then we have the following:*

(i) *If a is atomless, then $\chi(a) = 0$;*

(ii) *$a \in A_1$ if and only if $\chi(a) = \mathbb{P}(a)$;*

(iii) *$\chi(a) \leq \chi(a_{\mathrm{at}}) \leq \mathbb{P}(b)$, where b is an atom of largest measure in a if it exists and $b = 0$ otherwise;*

(iv) *$\varphi^{-1}(0) = A_1$, A_1 is closed, and $d(a, A_1) \leq \varphi(a)$;*

(v) *A_1 is 0-definable and it is defined by φ_1;*

(vi) *a is atomless if and only if $\theta(a) = 0$.*

Proof. Item (i) and (ii) follow from Corollary 4.12. Item (iii) follows from item (i), and the fact that for $a_{\mathrm{at}} = \bigvee_n b_n$, an atomic representation of a, then

$$(\mathbb{P}(b_1) + \mathbb{P}(b_3) + \dots) - (\mathbb{P}(b_2) + \mathbb{P}(b_4) + \dots) \leq \mathbb{P}(b_1).$$

For item (iv), observe that by item (ii) we have $\varphi^{-1}(0) = A_1$, respectively A_1 is closed. Observe that if a is atomless or zero, then $d(a, A_1) = \mathbb{P}(a)$ and $\varphi(a) = \mathbb{P}(a)$. If a is not atomless and nonzero, then $d(a, A_1) = \mathbb{P}(a) - \mathbb{P}(b)$, where b is an atom of largest measure in a, and $\mathbb{P}(a) - \mathbb{P}(b) \leq \varphi(a)$, by item (iii). Therefore for every a we have $d(a, A_1) \leq \varphi(a)$. Finally, by item (iv) and Proposition 2.33, we obtain item (v), that is A_1 is 0-definable, and similarly as in the proof of Proposition 2.33 (where with notations from its proof, $\alpha : [0,1] \to [0,1]$ is taken to be the identity) we obtain that φ_1 defines A_1. □

Using Remark 2.34 and that A_1 is 0-definable, we can prove that for every $n \geq 2$

$$A_n^{\boldsymbol{M}} = A_n = \{a_1 + a_2 + \cdots + a_n : a_m \in A_1\},$$

is also 0-definable. First, $A_1^n \subseteq M^n$ is defined by $\max(\varphi_1(x_1), \dots, \varphi_1(x_n))$ where $x_1, \dots, x_n$ are single variables and then A_n is defined by

$$\varphi_n(x) = \inf_{(x_1,\dots,x_n) \in A_1^n} d(x, x_1 \vee x_2 \vee \cdots \vee x_n).$$

Remark 4.15. *Let $a \in M$ and let $a_{\mathrm{at}} = \bigvee a_n$ be an atomic representation of a. Then for each $n > 0$ we have*

$$\varphi_n(a) = d(a, A_n) = \mathbb{P}(a) - \mathbb{P}(a_1) - \cdots - \mathbb{P}(a_n).$$

It follows that, $\mathbb{P}(a) \geq \varphi_1(a) \geq \varphi_2(a) \geq \dots$.

With notations from 4.15 observe that $\mathbb{P}(a_n)$ is *the n'th largest measure of an atom in a*. We denote $\mathrm{at}_n^{\boldsymbol{M}}(a) = \mathbb{P}(a_n)$. In fact, this is $\{a\}$-definable,

(i) $\mathrm{at}_1^{\boldsymbol{M}}(a) = \mathbb{P}(a) - \varphi_1^{\boldsymbol{M}}(a)$ and,

(ii) for each $n > 1$, $\mathrm{at}_n^{\boldsymbol{M}}(a) = \varphi_{n-1}^{\boldsymbol{M}}(a) - \varphi_n^{\boldsymbol{M}}(a)$.

Definition 4.16. Let $\mathcal{L}_{Pr^*}$ be the expansion of $\mathcal{L}_{Pr}$ by relation symbols $\{\mathrm{at}_1, \mathrm{at}_2, \dots\}$ and Pr^* be the $\mathcal{L}_{Pr^*}$-theory axiomatized by Pr and

$$\{\sup_x |\mathrm{at}_1(x) - (\mathbb{P}(x) - \varphi_1(x))| = 0\} \cup \{\sup_x |\mathrm{at}_n(x) - (\varphi_{n-1}(x) - \varphi_n(x))| = 0 \ : n \geq 2\}.$$

Remark 4.17. *Note that since the only function symbols of $\mathcal{L}_{Pr}$ are $0, 1, (\cdot)^c, \vee$ and $\wedge$, the substructures of $\boldsymbol{M}$ are the boolean σ-subalgebras of $\boldsymbol{M}$. In particular, finite boolean subalgebras of $\boldsymbol{M}$ are substructures of $\boldsymbol{M}$.*

Proposition 4.18. *Let $\boldsymbol{M}, \boldsymbol{N} \models Pr^*$, and let Γ denote the set of all partial isomorphisms between $\boldsymbol{M}$ and $\boldsymbol{N}$ whose domains are finite boolean subalgebras of $\boldsymbol{M}$. If Γ is nonempty, then it is a back-and-forth system.*

Proof. Let $\gamma \in \Gamma$, with $\gamma : A \to B$, and let $a \in M \setminus \mathrm{dom}(\gamma)$. We need to prove there exists an extension $\gamma' \in \Gamma$ of γ with $a \in \mathrm{dom}(\gamma')$. Let $\{a_1, \ldots, a_n\} \subseteq A, \{b_1, \ldots, b_n\} \subseteq B$ be the set of atoms of A and of B, such that $\gamma(a_m) = b_m$ for each m. Since, for each $m \geq 1$ and $1 \leq i \leq n$, we have $\mathrm{at}_m^{\boldsymbol{M}}(a_i) = \mathrm{at}_m^{\boldsymbol{N}}(b_i)$, we obtain a bijective measure-preserving correspondence between the atoms of a_i and b_i. Moreover, $\mathbb{P}^{\boldsymbol{M}}((a_i)_{\mathrm{at}}) = \mathbb{P}^{\boldsymbol{N}}((b_i)_{\mathrm{at}})$ and $\mathbb{P}^{\boldsymbol{M}}((a_i)_{\mathrm{al}}) = \mathbb{P}^{\boldsymbol{N}}((b_i)_{\mathrm{al}})$.

For each $1 \leq i \leq n$ let $u_i = a \wedge a_i$. Let $(u_i)_{\mathrm{at}} = \bigvee_m (u_i)_m$ be an atomic representation of u_i. Then $(u_i)_{\mathrm{al}} \leq (a_i)_{\mathrm{al}}$. By the correspondence mentioned in the previous paragraph and Corollary 4.12, results that there exists $v_i \leq b_i$ such that for the atomic representation $(v_i)_{\mathrm{at}} = \bigvee_m (v_i)_m$, we have $\mathbb{P}((v_i)_m) = \mathbb{P}((u_i)_m)$ and $\mathbb{P}((v_i)_{\mathrm{al}}) = \mathbb{P}((u_i)_{\mathrm{al}})$.

Finally, it remains to observe that there exists a unique partial isomorphism γ' from the subalgebra of $\boldsymbol{M}$ generated by A and u onto the subalgebra of $\boldsymbol{N}$ generated by B and v, extending γ, such that $\gamma'(u) = v$ and $\gamma'(u_i) = v_i$ for each $1 \leq i \leq n$. □

Using this result we can prove Pr^* has quantifier elimination.

Theorem 4.19. *The theory Pr^* has QE.*

Proof. We apply Theorem 2.69. Let φ be an $\mathcal{L}(x)$-formula, $\boldsymbol{M}, \boldsymbol{N} \models Pr^*$, A a common substructure of $\boldsymbol{M}$ and $\boldsymbol{N}$, and $a \in A_x$. We need to prove that $\varphi^{\boldsymbol{M}}(a) = \varphi^{\boldsymbol{N}}(a)$. We can assume A is a finite boolean subalgebra. Let Γ denote the set of all partial isomorphisms between $\boldsymbol{M}$ and $\boldsymbol{N}$ whose domains are finite boolean subalgebras of $\boldsymbol{M}$. Clearly, $id : A \to A$ is in Γ. By Proposition 4.18, Γ is a back-and-forth system, and by Theorem 2.65, each $\gamma \in \Gamma$ is an elementary map. Then, $id : A \to A$ is an elementary map and we obtain $\varphi^{\boldsymbol{M}}(a) = \varphi^{\boldsymbol{N}}(a)$. □

Using notations from Proposition 4.18, note that for $\boldsymbol{M}, \boldsymbol{N} \models Pr^*$, Γ is a back-and-forth system if and only if the map

$$\gamma : \{0^{\boldsymbol{M}}, 1^{\boldsymbol{M}}\} \to \{0^{\boldsymbol{N}}, 1^{\boldsymbol{N}}\}, \text{with } 0^{\boldsymbol{M}} \to 0^{\boldsymbol{N}}, 1^{\boldsymbol{M}} \to 1^{\boldsymbol{N}},$$

is a partial isomorphism from $\boldsymbol{M}$ to $\boldsymbol{N}$. Furthermore, by Corollary 2.66, if Γ is a back-and-forth system, we have $\boldsymbol{M} \equiv \boldsymbol{N}$. Conversely, if $\boldsymbol{M} \equiv \boldsymbol{N}$ then γ is a partial isomorphism, and respectively Γ is a back-and-forth system. We denote $\mathrm{at}^{\boldsymbol{M}} := (\mathrm{at}^{\boldsymbol{M}}(1))_{n \geq 1}$. Summarizing, we obtain:

Proposition 4.20. *[BH23, Corollary 4.16] For $\boldsymbol{M}, \boldsymbol{N} \models Pr$, we have $\boldsymbol{M} \equiv \boldsymbol{N}$ if and only if $\mathrm{at}^{\boldsymbol{M}} = \mathrm{at}^{\boldsymbol{N}}$. Additionally, for an extension T of Pr (in $\mathcal{L}_{Pr}$) the following are equivalent:*

(i) T is complete;

(ii) there exists a decreasing sequence of reals $(t_n)_{n \geq 1}$ satisfying $0 \leq t_n \leq 1$ for each $n \geq 1$, and $\sum_{n \geq 1} t_n \leq 1$, such that T is axiomatized by

$$Pr \cup \{\mathbb{P}(1) - \varphi_1(1) = t_1\} \cup \{\varphi_{n-1}(1) - \varphi_n(1) = t_n : n \geq 2\}.$$

Remark 4.21. *In Proposition 4.20, every model* $\boldsymbol{M} \models T$ *has its atomic part generated by a sequence of distinct elements* $(a_n)_{n \geq 1}$*, where each* a_n *is an atom or a zero, satisfying* $\mathbb{P}(a_n) = t_n$ *for every* $n \geq 1$*, and* $\boldsymbol{M}$ *has an atomless part of measure* $1 - \sum_{n \geq 1} t_n \leq 1$*. If* $1 - \sum_{n \geq 1} t_n = 0$ *then* $\boldsymbol{M}$ *has zero as its atomless part, and thus* T *has an unique model up to isomorphism.*

Remark 4.22. *Let* $\boldsymbol{M} \models Pr^*$ *and let* $\boldsymbol{N} \succcurlyeq \boldsymbol{M}$ *be* $\aleph_0$*-saturated. Then* Γ *is a back-and-forth system between* $\boldsymbol{M}$ *and* $\boldsymbol{N}$*, hence* $\boldsymbol{M}$ *is also* $\aleph_0$*-saturated. Therefore every model of* Pr^**, and consequently of* Pr*, is* $\aleph_0$*-saturated.*

We call a theory T *separably categorical* if it has only one model of countable size (has a countable dense set) up to isomorphism.

Proposition 4.23. *[BH23, Corollary 4.17] Every completion* T *of* Pr *is separably categorical, and the unique separable model of* T *is strongly* $\aleph_0$*-homogeneous.*

Proof. Let $\boldsymbol{M}, \boldsymbol{N} \models T$ be two separable $\mathcal{L}_{Pr}$-structures. Then using the fact that $\boldsymbol{M}, \boldsymbol{N}$ are $\aleph_0$-saturated, and have countable dense sets, then by a standard back-and-forth argument in ω steps we obtain an elementary map $f : A \to B$, where A is dense in $\boldsymbol{M}$ and B is dense in $\boldsymbol{N}$. Clearly, f can be extended to an isomorphism between $\boldsymbol{M}$ and $\boldsymbol{N}$.

Finally, the unique separable model is strongly $\aleph_0$-homogeneous as a consequence of the fact that it is $\aleph_0$-saturated (one can extend any finite elementary map $\boldsymbol{M} \to \boldsymbol{M}$ by the same construction as in the previous paragraph). $\square$

We end this section with a description of the separable models of Pr and the 0-definable elements in models of Pr.

Definition 4.24. A *measure algebra* is a tuple

$$\boldsymbol{B} = (B, 0, 1, \cdot^c, \wedge, \vee, \mu)$$

where $(B, 0, 1, \cdot^c, \wedge, \vee)$ is a σ-complete boolean algebra (every countable set of elements has a supremum and infimum) and μ is a strictly positive measure on B (for every $a \in B, 0 < a \implies 0 < \mu(a)$).

Let T be a completion of Pr and let $\boldsymbol{M}$ be its unique separable model. Then $(1^{\boldsymbol{M}})^0 = \bigvee_{n \geq 1} a_n$, where each a_n is an atom or zero, satisfying $\mathbb{P}^{\boldsymbol{M}}(a_n) = t_n$, where t_n is the unique real such that

$$T \models \mathbb{P}(1) \dot{-} \varphi_1(1) = t_1 \text{ and } T \models \varphi_{n-1}(1) \dot{-} \varphi_n(1) = t_n \text{ for each } n \geq 2.$$

While the atomless part $1^{\boldsymbol{M}} \backslash (1^{\boldsymbol{M}})^0$, as a structure with the σ-algebra contained in $1^{\boldsymbol{M}} \backslash (1^{\boldsymbol{M}})^0$ and σ-measure $\mathbb{P}^{\boldsymbol{M}}$, is isomorphic to the measure algebra associated to the Lebesgue measure on the interval $[0, 1 - \sum_{n \geq 1} t_n]$.

Proposition 4.25. *Let* $\boldsymbol{M} \models Pr$ *and* $C \subseteq M$*. Then* $\mathrm{dcl}(C)$ *is a* σ*-algebra. Moreover, for every* $a \in \mathrm{dcl}(C)$ *we have* $a_{at}, a_{al} \in \mathrm{dcl}(C)$*.*

Proof. By Proposition 2.55 we obtain that $\mathrm{dcl}(C)$ is closed under $\wedge, \vee, (\cdot)^c$, and contains $0^{\boldsymbol{M}}, 1^{\boldsymbol{M}}$. Thus, $\mathrm{dcl}(C)$ is a boolean algebra. Let $c_0 \leq c_1 \leq c_2 \leq c_3 \leq \ldots$ be a sequence of C-definable elements. Then for every $a \in M$ we have $d^{\boldsymbol{M}}(a, \vee c_n) = \lim_n d^{\boldsymbol{M}}(a, c_n)$. Moreover the map $d^{\boldsymbol{M}}(\cdot, \vee c_n)$ is the uniform limit of the definable predicates $d^{\boldsymbol{M}}(\cdot, c_n)$, hence $d^{\boldsymbol{M}}(\cdot, \vee c_n)$ is a definable predicate, and we get that $\vee c_n$ is C-definable. Therefore $\mathrm{dcl}(C)$ is a σ-algebra.

For the second part, let $\mathbb{M} \succcurlyeq \boldsymbol{M}$ be a monster model. Then, for every $a \in \mathrm{dcl}(C)$ and $h \in \mathrm{Aut}(\mathbb{M}|C)$ we have $h(a) = a$, by $\mathcal{C}$-definability of a. Moreover, since h is an automorphism, it follows that $h(a_{\mathrm{at}}) = a_{\mathrm{at}}$ and $h(a_{\mathrm{al}}) = a_{\mathrm{al}}$. Applying Theorem 2.62 we obtain that a_{at} and a_{al} are C-definable. $\square$

4.3 Stability of APA

The aim of this section is to prove that the theory of atomless probability algebras is $\aleph_0$-stable. We begin with a review of some concepts from probability.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{D} \subseteq \mathcal{F}$ be a σ-subalgebra. For each $n \geq 1$, we use the standard notation

$$L^n(\Omega, \mathcal{F}, \mathbb{P}) = \{f : \Omega \to \mathbb{R} : \ f \text{ is } \mathcal{F}\text{-measurable and } \|f\|_n = \left(\int_\Omega |f|^n d\mathbb{P}\right)^{\frac{1}{n}} < \infty\}/\sim,$$

where $\sim$ is an equivalence relation where $f \sim g$ if and only if $f = g$ almost surely, that is, they differ on a set of at most measure zero. In what follows, we will refer to elements of $L^n(\Omega, \mathcal{F}, \mathbb{P})$ by representative functions, with the understanding that any statement or operation involving such functions is invariant under almost sure equivalence.

Proposition 4.26. *[Dud18, Theorem 10.1.1] For any function $f \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ there exists a function $g \in L^1(\Omega, \mathcal{D}, \mathbb{P})$, such that for any $A \in \mathcal{D}$ we have*

$$\int_A f \ d\mathbb{P} = \int_A g \ d\mathbb{P}.$$

We call g a conditional expectation of f with respect to $\mathcal{D}$. Any two conditional expectations of f with respect to $\mathcal{D}$ are equal up to a measure zero set. Therefore, we can speak of the conditional expectation of f with respect to $\mathcal{D}$, denoted by $\mathbb{E}(f|\mathcal{D})$.

In other words, $\mathbb{E}(f|\mathcal{D})$ is the best approximation of f by a $\mathcal{D}$-measurable L^1-function. Let $\boldsymbol{M} = (M, 0, 1, \cdot^c, \vee, \wedge, \mathbb{P}^{\boldsymbol{M}})$ be the probability algebra corresponding to $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathcal{C} \subseteq M$ be a σ-subalgebra. By Proposition 4.2, there exists a σ-subalgebra $\mathcal{D} \subseteq \mathcal{F}$ containing all the null sets, such that $\widehat{\mathcal{D}} = \mathcal{C}$. Then for any $a \in M$ we define

$$\mathbb{P}^{\boldsymbol{M}}(a|\mathcal{C}) := \mathbb{E}(\chi_A|\mathcal{D}),$$

where $\widehat{A} = a$. This is well-defined, that is, if $\widehat{A} = \widehat{B} = a$, then $\chi_A \sim \chi_B$, and therefore $\mathbb{E}(\chi_A|\mathcal{D}) = \mathbb{E}(\chi_B|\mathcal{D})$. Since for any $c \in \mathcal{C}$ and $\widehat{C} = c$ we have

$$\int_C \mathbb{P}^{\boldsymbol{M}}(a|\mathcal{C}) \ d\mathbb{P} = \mathbb{P}^{\boldsymbol{M}}(a \wedge c),$$

by Proposition 4.26, $\mathbb{P}^{\boldsymbol{M}}(a|\mathcal{C})$ is uniquely determined by the values $\mathbb{P}^{\boldsymbol{M}}(a \wedge c)$ for $c \in \mathcal{C}$. Therefore, using the same notations, we obtain:

Proposition 4.27. *Let $a, b \in M$. Then $\mathbb{P}^{\boldsymbol{M}}(a|\mathcal{C}) = \mathbb{P}^{\boldsymbol{M}}(b|\mathcal{C})$ if and only if $\mathbb{P}^{\boldsymbol{M}}(a \wedge c) = \mathbb{P}^{\boldsymbol{M}}(b \wedge c)$ for all $c \in \mathcal{C}$. Furthermore, if $C \subseteq \mathcal{C}$ is a boolean algebra which generates $\mathcal{C}$, then $\mathbb{P}^{\boldsymbol{M}}(a|\mathcal{C}) = \mathbb{P}^{\boldsymbol{M}}(b|\mathcal{C})$ if and only if $\mathbb{P}^{\boldsymbol{M}}(a \wedge c) = \mathbb{P}^{\boldsymbol{M}}(b \wedge c)$ for all $c \in C$.*

Proof. The first part of the statement is clear by the argument in the previous paragraph. For the furthermore part, let $\mathcal{S} = \{s \in M : \mathbb{P}^{\boldsymbol{M}}(a \wedge s) = \mathbb{P}^{\boldsymbol{M}}(b \wedge s)\}$. Then, by assumption, $C \subseteq \mathcal{S}$, and by properties of $\mathbb{P}^{\boldsymbol{M}}$, it follows that $\mathcal{S}$ is a monotone class, that is, for any $s_1 \leq s_2 \leq \dots$ we have $\vee_n s_n \in \mathcal{S}$ and for all $s_1 \geq s_2 \geq \dots$ we have $\wedge s_n \in \mathcal{S}$. By the monotone class theorem [Dud18, Theorem 4.4.2], we obtain that $\mathcal{C} \subseteq \mathcal{S}$. $\square$

Remark 4.28. *For $a, b \in M$ by definition we have $\mathbb{P}^{\boldsymbol{M}}(a|\mathcal{C}) \leq \mathbb{P}^{\boldsymbol{M}}(b|\mathcal{C})$ if and only if $\mathbb{E}(\chi_A|\mathcal{D}) \leq \mathbb{E}(\chi_B|\mathcal{D})$, for some $\widehat{A} = a$ and $\widehat{B} = b$. Hence, it is easy to see that $\mathbb{P}^{\boldsymbol{M}}(a|\mathcal{C}) \leq \mathbb{P}^{\boldsymbol{M}}(b|\mathcal{C})$ if and only if $\mathbb{P}^{\boldsymbol{M}}(a \wedge c) \leq \mathbb{P}^{\boldsymbol{M}}(b \wedge c)$ for every $c \in \mathcal{C}$.*

Observe that $\mathbb{P}^{\boldsymbol{M}}(a|\mathcal{C})$ depends on the probability algebra $(\Omega, \mathcal{F}, \mathbb{P})$. However, we are more interested when $\mathbb{P}^{\boldsymbol{M}}(a|\mathcal{C}) \leq \mathbb{P}^{\boldsymbol{M}}(b|\mathcal{C})$ or $\mathbb{P}^{\boldsymbol{M}}(a|\mathcal{C}) = \mathbb{P}^{\boldsymbol{M}}(b|\mathcal{C})$. By Proposition 4.27 and Remark 4.28, these comparative relations remain invariant under any choice of probability space $(\Omega, \mathcal{F}, \mathbb{P})$ whose probability algebra is $\boldsymbol{M}$. Hence, we record:

Proposition 4.29. *Let $a, b \in M$. Then the relations $\mathbb{P}^{\boldsymbol{M}}(a|\mathcal{C}) \leq \mathbb{P}^{\boldsymbol{M}}(b|\mathcal{C})$ and $\mathbb{P}^{\boldsymbol{M}}(a|\mathcal{C}) = \mathbb{P}^{\boldsymbol{M}}(b|\mathcal{C})$ are independent of the choice of probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with probability algebra $\boldsymbol{M}$.*

Conditional expectation can also be interpreted as a projection in a Hilbert space. We will assume the reader knows the basic properties of Hilbert spaces, as in Chapter 5 of [Dud18]. The space $L^2(\Omega, \mathcal{F}, \mathbb{P})$ with the inner product $(f, g) = \int_\Omega fg \, d\mathbb{P}$ is a Hilbert space with norm $\|f\|_2 = \sqrt{(f, f)} = \sqrt{\int_\Omega f^2}$. The subspace $L^2(\Omega, \mathcal{D}, \mathbb{P})$ is a closed subspace of $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Then the projection of $f \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ onto $L^2(\Omega, \mathcal{D}, \mathbb{P})$ is $\mathbb{E}(f|\mathcal{D})$ ([Dur19, Theorem 4.1.15]).

Since $f = \mathbb{E}(f|\mathcal{D}) + (f - \mathbb{E}(f|\mathcal{D}))$ and $\mathbb{E}(f|\mathcal{D}) \perp (f - \mathbb{E}(f|\mathcal{D}))$, then $\|f\|_2^2 = \|\mathbb{E}(f|\mathcal{D})\|_2^2 + \|f - \mathbb{E}(f|\mathcal{D})\|_2^2$, in particular $\|f\|_2 \geq \|\mathbb{E}(f|\mathcal{D})\|_2$, in other words the projected vector always has smaller length. Let $\mathcal{H} \subseteq \mathcal{E} \subseteq \mathcal{D}$ be a chain of σ-subalgebras. Then by basic conditional expectation properties [Dur19, Theorem 4.1.13], we have that $\mathbb{E}(f|\mathcal{E}) = \mathbb{E}(\mathbb{E}(f|\mathcal{D})|\mathcal{E})$. Furthermore, since $\mathbb{E}(f|\mathcal{D}) - \mathbb{E}(f|\mathcal{E}) \perp \mathbb{E}(f|\mathcal{E})$ we have

$$\|\mathbb{E}(f|\mathcal{D}) - \mathbb{E}(f|\mathcal{E})\|_2^2 = \|\mathbb{E}(f|\mathcal{D})\|_2^2 - \|\mathbb{E}(f|\mathcal{E})\|_2^2,$$

similarly

$$\|\mathbb{E}(f|\mathcal{D}) - \mathbb{E}(f|\mathcal{H})\|_2^2 = \|\mathbb{E}(f|\mathcal{D})\|_2^2 - \|\mathbb{E}(f|\mathcal{H})\|_2^2,$$

and coupled with the fact that $\|\mathbb{E}(f|\mathcal{H})\|_2 \leq \|\mathbb{E}(f|\mathcal{E})\|_2$, we get

$$\|\mathbb{E}(f|\mathcal{D}) - \mathbb{E}(f|\mathcal{E})\|_2 \leq \|\mathbb{E}(f|\mathcal{D}) - \mathbb{E}(f|\mathcal{H})\|_2.$$

Note that by the Cauchy-Schwarz inequality applied for f and the constant function 1, we have $\|f\| \leq \|f\|_2$. Therefore we get,

$$\|\mathbb{E}(f|\mathcal{D}) - \mathbb{E}(f|\mathcal{E})\| \leq \|\mathbb{E}(f|\mathcal{D}) - \mathbb{E}(f|\mathcal{E})\|_2 \leq \|\mathbb{E}(f|\mathcal{D}) - \mathbb{E}(f|\mathcal{H})\|_2.$$

Now, by translating to the probability algebra $\boldsymbol{M}$ and taking $f = \chi_A$ for $\widehat{A} = a$ and considering $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{C}$ a chain of σ-subalgebras we obtain [BH23, p. 6],

$$\|\mathbb{P}^{\boldsymbol{M}}(a|\mathcal{C}) - \mathbb{P}^{\boldsymbol{M}}(a|\mathcal{B})\| \leq \|\mathbb{P}^{\boldsymbol{M}}(a|\mathcal{C}) - \mathbb{P}^{\boldsymbol{M}}(a|\mathcal{B})\|_2 \leq \|\mathbb{P}^{\boldsymbol{M}}(a|\mathcal{C}) - \mathbb{P}^{\boldsymbol{M}}(a|\mathcal{A})\|_2.$$

Remark 4.30. *Let $f \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Then, we already know that $\|\mathbb{E}(f|\mathcal{D})\|_2 \leq \|f\|_2$, as $\mathbb{E}(\cdot|\mathcal{D})$ is a projection and hence a contraction. By definition, we know that $\int_\Omega f d\mathbb{P} = \int_\Omega \mathbb{E}(f|\mathcal{D}) d\mathbb{P}$, but it is not generally true that $\int_\Omega |f| d\mathbb{P} = \int_\Omega |\mathbb{E}(f|\mathcal{D})| d\mathbb{P}$. In fact we have $\|\mathbb{E}(f|\mathcal{D})\|_1 \leq \|f\|_1$. Which can be seen from the following*

$$\|\mathbb{E}(f|\mathcal{D})\|_1 \leq \int_\Omega \mathbb{E}(f|\mathcal{D}) \cdot \operatorname{sgn}(\mathbb{E}(f|\mathcal{D}))\ d\mathbb{P} = \int_\Omega f \cdot \operatorname{sgn}(\mathbb{E}(f|\mathcal{D}))\ d\mathbb{P} \leq \int_\Omega |f| d\mathbb{P} = \|f\|_1,$$

where sgn is the sign function, and the second equality is because of the fact that for any $g \in L^1(\Omega, \mathcal{D}, \mathbb{P})$ we have $\mathbb{E}(fg|\mathcal{D}) = \mathbb{E}(f|\mathcal{D})g$. Therefore $\mathbb{E}(\cdot|\mathcal{D})$ is a contraction in L^1 as well.

Remark 4.31. *Let $f \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Then it is easy to see that if $\operatorname{im}(f) \subseteq [a, b]$, for some $a, b \in \mathbb{R}$, then $\operatorname{im}(\mathbb{E}(f|\mathcal{D})) \subseteq [a, b]$. Furthermore, if we restrict to a set $A \subseteq \mathcal{D}$, then similarly, if $\operatorname{im}(f|_A) \subseteq [a, b]$ then $\operatorname{im}(\mathbb{E}(f|\mathcal{D})|_A) \subseteq [a, b]$.*

Finally we can state a useful lemma.

Lemma 4.32. *Let $f \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, with $\operatorname{im}(f) \subseteq [0, 1]$. Then for every $n > 0$ there are $A_1, \ldots, A_n \in \mathcal{F}$ partitioning Ω such that for any σ-subalgebra $\mathcal{D}$ containing $A_1, \ldots, A_n$, we have*

$$\|f - \mathbb{E}(f|\mathcal{D})\|_1 \leq 1/n.$$

In particular $\|f - \mathbb{E}(f|\mathrm{BA}(A_1, \ldots, A_n))\|_1 \leq 1/n$.

Proof. Fix $n > 0$, and for each $0 \leq m < n$, let $A_m = f^{-1}([\frac{m}{n}, \frac{m+1}{n}])$. Then, by Remark 4.31, we obtain that $|f - \mathbb{E}(f|\mathcal{A})| \leq 1/n$, almost everywhere. Consequently, $\|f - \mathbb{E}(f|\mathcal{A})\|_1 \leq 1/n$. □

By translating to the probability algebra $\boldsymbol{M}$ and choosing $f = \mathbb{E}(\chi_A|\mathcal{D})$, we obtain:

Lemma 4.33. *[BH23, Lemma 2.7] Let $a \in M$. Then for every $n > 0$, there exists a paritition $\{a_1, \ldots, a_m\}$ of $1^{\boldsymbol{M}}$, such that for any σ-subalgebra $\mathcal{A} \subseteq \mathcal{C}$ containing the elements $a_1, \ldots, a_m$ we have*

$$\|\mathbb{P}^{\boldsymbol{M}}(a|\mathcal{C}) - \mathbb{P}^{\boldsymbol{M}}(a|\mathcal{A})\|_1 \leq 1/n.$$

In particular $\|\mathbb{P}^{\boldsymbol{M}}(a|\mathcal{C}) - \mathbb{P}^{\boldsymbol{M}}(a|\mathrm{BA}(a_1, \ldots, a_m))\|_1 \leq 1/n$.

Proof. This follows from Lemma 4.32 for $f = \mathbb{E}(\chi_A|\mathcal{D})$, with $\widehat{A} = a$, and observing that for $\mathcal{B} \subseteq \mathcal{D}$, such that $\widehat{\mathcal{B}} = \widehat{\mathcal{A}}$, we have $\mathbb{E}(\chi_A|\mathcal{B}) = \mathbb{E}(\mathbb{E}(\chi_A|\mathcal{D})|\mathcal{B})$. □

We end this section with a construction of a probability algebra.

Let $\boldsymbol{M}_n = (M_n, 0^{\boldsymbol{M}_n}, 1^{\boldsymbol{M}_n}, ((\cdot)^c)^{\boldsymbol{M}_n}, \wedge^{\boldsymbol{M}_n}, \vee^{\boldsymbol{M}_n}, \mu^{\boldsymbol{M}_n})$, be measure algebras for each n. Assume that $\sum_n \mu^{\boldsymbol{M}_n}(1^{\boldsymbol{M}_n}) = 1$. Then we define the sum of $(\boldsymbol{M}_n)$ as the probability algebra

$$\sum_n \boldsymbol{M}_n = \boldsymbol{S} = \left(\prod_n M_n, , 0^{\boldsymbol{S}}, 1^{\boldsymbol{S}}, ((\cdot)^c)^{\boldsymbol{S}}, \wedge^{\boldsymbol{S}}, \vee^{\boldsymbol{S}}, \mathbb{P}^{\boldsymbol{S}}\right),$$

where for $\prod_n a_n, \prod_n b_n \in \prod_n M_n$ we have

(i) $0^{\boldsymbol{S}} = \prod_n 0^{\boldsymbol{M}_n}, 1^{\boldsymbol{S}} = \prod_n 1^{\boldsymbol{M}_n}$;

(ii) $\prod_n a_n \wedge^{\boldsymbol{S}} \prod_n b_n = \prod_n a_n \wedge^{\boldsymbol{M}_n} b_n$, and $\prod_n a_n \vee^{\boldsymbol{S}} \prod_n b_n = \prod_n a_n \vee^{\boldsymbol{M}_n} b_n$;

(iii) $((\prod_n a_n)^c)^{\boldsymbol{S}} = \prod_n (a_n^c)^{\boldsymbol{M}_n}$;

(iv) and finally $\mathbb{P}^{\boldsymbol{S}}(\prod_n a_n) = \sum_n \mu^{\boldsymbol{M}_n}(a_n)$.

Atomless Probability Algebras

Let $\boldsymbol{M} \models Pr$. By Proposition 4.14 we know that $\boldsymbol{M}$ is an atomless probability algebra if and only if $\theta^{\boldsymbol{M}}(1^{\boldsymbol{M}}) = 0$. This condition is also equivalent with

$$\boldsymbol{M} \models \sup_x \inf_y |\mathbb{P}(x \wedge y) - \mathbb{P}(x \wedge y^c)| = 0.$$

We call $\sup_x \inf_y |\mathbb{P}(x \wedge y) - \mathbb{P}(x \wedge y^c)| = 0$ the *atomless axiom.*

Definition 4.34. We denote by APA the theory axiomatized by Pr together with the atomless axiom.

It is clear that the models of APA are exactly the atomless probability algebras.

Definition 4.35. Just as in classical model theory, we call a theory T *model complete* if for any $\boldsymbol{M}, \boldsymbol{N} \models T$ and $\boldsymbol{M} \subseteq \boldsymbol{N}$ we have $\boldsymbol{M} \preccurlyeq \boldsymbol{N}$. A theory T^* is called a *model companion* of T if T^* is model complete, and every model of T embeds into in a model of T^*.

Theorem 4.36. *[BH23, Corollary 6.2] The theory APA has QE, is separably categorical, is complete, and the unique separable model of APA is strongly $\aleph_0$-homogeneous. Furthermore, APA is the model companion of Pr.*

Proof. Let APA^* be the extension by definitions of the theory APA where we add $\mathrm{at_n}$ as definable predicates for each $n \geq 1$. Since $Pr^* \subseteq APA^*$ and the former has QE, so does the latter. But, since $APA^* \models \sup_x |\mathrm{at}_n(x)| = 0$, it follows trivially that APA has QE.

Note that APA is axiomatized by Pr and $|\varphi_1(1)| = 0$. By Proposition 4.20, it follows that APA is a completion of Pr and by Proposition 4.23, we obtain that APA is separably categorical and its unique separable model is strongly $\aleph_0$-homogeneous.

Since APA has QE, it is model complete. Let $\boldsymbol{M} \models Pr$ and let $\bigvee_n a_n$ be the atomic representation of $1^{\boldsymbol{M}}$. For each n, let $\boldsymbol{A}_n$ be the measure algebra associated to the Lebesgue measure algebra on the interval $[0, \mathbb{P}^{\boldsymbol{M}}(a_n)]$. Let B be the measure algebra on $(1)^{\boldsymbol{M}}_{al}$ inherited from $\boldsymbol{M}$. Then the sum of B and $(A_n)_n$ is an atomless probability algebra, into which $\boldsymbol{M}$ clearly embeds. □

Throughout the end of this section let $\boldsymbol{M} \models APA$ and $x = (x_1, \dots, x_n)$ be a finite multivariable.

Definition 4.37. For $C \subseteq M$, we denote by $\mathrm{BA}(C)$ the boolean subalgebra of M generated by C and by $\sigma(C)$ the σ-subalgebra of M generated by C. For $a \in M$ we denote $a^1 := a$ and $a^{-1} := a$. For $a \in M_x$ and $s = (s_1, \dots, s_n) \in \{-1, 1\}^n$ we denote $a^s := (a_1)^{s_1} \wedge (a_2)^{s_2} \wedge \dots \wedge (a_n)^{s_n}$. Simiarly, for x, we define the $\mathcal{L}$-term x^s.

Note that for $a \in M_x$, $\{a^s : s \in \{-1,1\}^n\} \backslash \{0\}$ is the set of atoms of the boolean subalgebra of M generated by $a_1, \dots, a_n$.

Remark 4.38. *Let $a \in M_x$ and $C \subseteq M$. Then since APA has QE it follows that $\mathrm{tp}^{\boldsymbol{M}}(a|C)$ is determined by its quantifier-free conditions, and as a result by the atomic conditions included in $\mathrm{tp}^{\boldsymbol{M}}(a|C)$. The atomic formulas are of the form $d(t_1, t_2)$ or $\mathbb{P}(t_1)$, for some $\mathcal{L}$-terms t_1 and t_2. Since $Pr \models \sup_{xy} |d(x,y) - \mathbb{P}(x \triangle y)| = 0$, the type $\mathrm{tp}^{\boldsymbol{M}}(a|C)$ is uniquely determined by the conditions of the form $\mathbb{P}(x^s \wedge c) = (\mathbb{P}(x^s \wedge c))^{\mathrm{tp}(a|C)}$ for $s \in \{-1,1\}^n$ and $c \in \mathrm{BA}(C)$.*

Next we present a lemma characterizing when two types are equal.

Lemma 4.39. *[BH23, Lemma 6.4] Let $\boldsymbol{M} \models APA$, $C \subseteq M$ and $a, b \in M_x$. The following are equivalent:*

(i) $\mathrm{tp}(a|C) = \mathrm{tp}(b|C)$*;*

(ii) $\mathbb{P}^{\boldsymbol{M}}(a^s \wedge c) = \mathbb{P}^{\boldsymbol{M}}(b^s \wedge c)$ *for every* $s \in \{-1,1\}^n$ *and* $c \in \mathrm{BA}(C)$*;*

(iii) $\mathbb{P}^{\boldsymbol{M}}(a^s \wedge c) = \mathbb{P}^{\boldsymbol{M}}(b^s \wedge c)$ *for every* $s \in \{-1,1\}^n$ *and* $c \in \sigma(C)$*;*

(iv) $\mathbb{P}^{\boldsymbol{M}}(a^s|\sigma(C)) = \mathbb{P}^{\boldsymbol{M}}(b^s|\sigma(C))$ *for every* $s \in \{-1,1\}^n$.

Proof. The equivalence $(i) \Leftrightarrow (ii)$ is immediate by Remark 4.38. Suppose item (ii). We observe that, for any $s \in \{-1,1\}^n$ we have $\{c \in M : \mathbb{P}^{\boldsymbol{M}}(a^s \wedge c) = \mathbb{P}^{\boldsymbol{M}}(b^s \wedge c)\}$ is a monotone class. Hence, by the monotone class theorem [Dud18, Theorem 4.4.2], $\sigma(C) \subseteq \{c \in M : \mathbb{P}^{\boldsymbol{M}}(a^s \wedge c) = \mathbb{P}^{\boldsymbol{M}}(b^s \wedge c)\}$, proving (iii). The converse is trivial. The equivalence $(iii) \Leftrightarrow (iv)$ follows from Proposition 4.27. □

Remark 4.40. *[BH23, Remark 6.5] Let $a, b \in M$. Then Lemma 4.39 simplifies to the following equivalence:*

(i) $\mathrm{tp}(a|C) = \mathrm{tp}(b|C)$*;*

(ii) $\mathbb{P}^{\boldsymbol{M}}(a \wedge c) = \mathbb{P}^{\boldsymbol{M}}(b \wedge c)$ *for every* $c \in \sigma(C)$*;*

(iii) $\mathbb{P}^{\boldsymbol{M}}(a|\sigma(C)) = \mathbb{P}^{\boldsymbol{M}}(b|\sigma(C))$*.*

Note that if $a = (a_1, \dots, a_n) \in M_x$ *is a partition of* $1^{\boldsymbol{M}}$ *then for every* $s \in \{-1, 1\}^n$*, we have that* a^s *is equal to* a_m*, for some* $1 \leq m \leq n$*, or* $0^{\boldsymbol{M}}$*. Therefore if* $b = (b_1, \dots, b_n) \in M_x$ *is also a partition of* $1^{\boldsymbol{M}}$ *then* $\mathrm{tp}(a|\mathcal{C}) = \mathrm{tp}(b|\mathcal{C})$ *if and only if* $\mathbb{P}^{\boldsymbol{M}}(a_m \wedge c) = \mathbb{P}^{\boldsymbol{M}}(b_m \wedge c)$ *for every* $c \in \sigma(\mathcal{C})$ *and* $1 \leq m \leq n$*, which is equivalent to* $\mathbb{P}^{\boldsymbol{M}}(a_m|\sigma(C)) = \mathbb{P}^{\boldsymbol{M}}(b_m|\sigma(C))$ *for every* $1 \leq m \leq n$*, and by Lemma 4.39 also equivalent to* $\mathrm{tp}(a_1|\mathcal{C}) = \mathrm{tp}(b_1|\mathcal{C}), \dots, \mathrm{tp}(a_n|\mathcal{C}) = \mathrm{tp}(b_n|\mathcal{C})$*.*

As a consequence of the above remarks we obtain that if $a = (a_1, \dots, a_n) \in M_x$ *is a partition of* $1^{\boldsymbol{M}}$*, then* $\mathrm{tp}(a|\mathcal{C})$ *is uniquely determined by* $\mathrm{tp}(a_1|\mathcal{C}), \dots, \mathrm{tp}(a_n|\mathcal{C})$*, in the sense that if* $q(x)$ *is any* x*-type over* $\mathcal{C}$*, having as a realization a partition of* $1^{\boldsymbol{M}}$*, and containing* $\mathrm{tp}(a_1|\mathcal{C}), \dots, \mathrm{tp}(a_n|\mathcal{C})$ *then* $q(x) = p(x)$*.*

Throughout the end of this section let $\mathbb{M}$ be a monster model of APA, and let $(\Omega, \mathcal{F}, \mathbb{P})$ be an atomless probability space with associated probability algebra $\mathbb{M}$. The next theorem is the main result in proving $\aleph_0$-stability of APA. For completeness, and as it provides an opportunity to apply some of the concepts presented in the introductory chapters, we present the proof as is written in [BH23, Theorem 6.11]. We recall the distance between two types, see the paragraph after Definition 1.78.

Theorem 4.41. *[BH23, Theorem 6.11] Let* $C \subseteq \mathbb{M}$ *be a small parameter set, and let* $a, b \in \mathbb{M}_x$*,* $a = (a_1, \dots, a_n)$ *and* $b = (b_1, \dots, b_n)$ *be partitions of* $1^{\mathbb{M}}$*. Then*

$$d(\mathrm{tp}(a|C), \mathrm{tp}(b|C)) = \max_{1 \leq m \leq n} \|\mathbb{P}^{\mathbb{M}}(a_m|\sigma(C)) - \mathbb{P}^{\mathbb{M}}(b_m|\sigma(C))\|_1.$$

Furthermore, there exists $b' = (b'_1, \dots, b'_n)$ *such that* $\mathrm{tp}(b|C) = \mathrm{tp}(b'|C)$ *and for all* $1 \leq m \leq n$ *we have*

$$d(a_m, b'_m) = \|\mathbb{P}^{\mathbb{M}}(a_m|\sigma(C)) - \mathbb{P}^{\mathbb{M}}(b_m|\sigma(C))\|_1.$$

Proof. The proof proceeds in four steps.

Step 1. Let $p(x) = \mathrm{tp}(a|\mathcal{C})$ and $q(x) = \mathrm{tp}(b|\mathcal{C})$. By Lemma 4.39 for any realizations $a' \models p, b' \models q$ we have $\mathbb{P}^{\mathbb{M}}(a_m|\sigma(C)) = \mathbb{P}^{\mathbb{M}}(a'_m|\sigma(C))$ and $\mathbb{P}^{\mathbb{M}}(b_m|\sigma(C)) = \mathbb{P}^{\mathbb{M}}(b'_m|\sigma(C))$. We also observe that by linearity of conditional expectation for each $1 \leq m \leq n$ we have

$$\begin{aligned}\|\mathbb{P}^{\mathbb{M}}(a'_m|\sigma(C)) - \mathbb{P}^{\mathbb{M}}(b'_m|\sigma(C))\|_1 &= \|\mathbb{P}^{\mathbb{M}}(a'_m \backslash b'_m|\sigma(C)) - \mathbb{P}^{\mathbb{M}}(b'_m \backslash a'_m|\sigma(C))\|_1 \\ &\leq \|\mathbb{P}^{\mathbb{M}}(a'_m \backslash b'_m|\sigma(C))\|_1 + \|\mathbb{P}^{\mathbb{M}}(b'_m \backslash a'_m|\sigma(C))\|_1 \\ &= \mathbb{P}^{\mathbb{M}}(a'_m \triangle b'_m) = d^{\mathbb{M}}(a'_m, b'_m).\end{aligned}$$

By letting a' and b' vary over realizations of p and q we obtain that

$$\max_{1 \leq m \leq n} \|\mathbb{P}^{\mathbb{M}}(a_m|\sigma(C)) - \mathbb{P}^{\mathbb{M}}(b_m|\sigma(C))\|_1 \leq d(p, q).$$

Therefore, it's enough to prove the reverse inequality,

$$\max_{1\leq m\leq n} \|\mathbb{P}^{\mathbb{M}}(a_m|\sigma(C)) - \mathbb{P}^{\mathbb{M}}(b_m|\sigma(C))\|_1 \leq d(p,q).$$

Step 2. We assume $C = \emptyset$. Once fixed p, q as in step 1 we can vary $a \models p$ and $b \models q$. In particular, since $\mathbb{M}$ is a monster model we can choose $a, b \in M_x$ such that $d^{\mathbb{M}}(a,b) = d(p,q)$. Let

$$I = \{m : \mathbb{P}^{\mathbb{M}}(a_m) \geq \mathbb{P}^{\mathbb{M}}(b_m)\}, \text{ and } J = \{m : \mathbb{P}^{\mathbb{M}}(a_m) < \mathbb{P}^{\mathbb{M}}(b_m)\}.$$

Since $\mathbb{M}$ is an atomless probability algebra, using Corollary 4.12, we can find a partition $b' = (b'_1, \dots, b'_n)$ of $1^{\mathbb{M}}$ such that $\mathbb{P}^{\mathbb{M}}(b'_m) = \mathbb{P}^{\mathbb{M}}(b_m)$ for every m, with the additional property that $a_m \geq b'_m$ for every $m \in I$ and $a_m < b'_m$ for every $m \in J$. Observe that $b' \models q$ and that for every m we have

$$d^{\mathbb{M}}(a_m, b'_m) \leq d^{\mathbb{M}}(a_m, b_m) \text{ and } \|\mathbb{P}^{\mathbb{M}}(a_m|\sigma(\emptyset)) - \mathbb{P}^{\mathbb{M}}(b'_m|\sigma(\emptyset))\|_1 = d^{\mathbb{M}}(a_m, b'_m).$$

Therefore

$$\max_{1\leq m\leq n} \|\mathbb{P}^{\mathbb{M}}(a_m|\sigma(\emptyset)) - \mathbb{P}^{\mathbb{M}}(b'_m|\sigma(\emptyset))\|_1 \leq d^{\mathbb{M}}(a,b') = d(p,q)$$

which also proves the furthermore part.

Step 3. We assume C is finite and as in step 2 let $a \models p, b \models q$ such that $d^{\mathbb{M}}(a,b) = d(p,q)$. In this case we can also assume $C = \sigma(C)$ is a finite boolean subalgebra of $\mathbb{M}$. Let $\{c_1, \dots, c_k\}$ be the atoms of C. Then, for each m we have that

$$d^{\mathbb{M}}(a_m, b_m) = d^{\mathbb{M}}(a_m \wedge c_1, b_m \wedge c_1) + \cdots + d^{\mathbb{M}}(a_m \wedge c_k, b_m \wedge c_k),$$

and

$$\|\mathbb{P}^{\mathbb{M}}(a_m|C) - \mathbb{P}^{\mathbb{M}}(b_m|C)\|_1 = |\mathbb{P}^{\mathbb{M}}(a_m \wedge c_1) - \mathbb{P}^{\mathbb{M}}(b_m \wedge c_1)| + \cdots + |\mathbb{P}^{\mathbb{M}}(a_m \wedge c_k) - \mathbb{P}^{\mathbb{M}}(b_m \wedge c_k)|$$

Therefore it suffices to find a realization $b' \models q$ such that $d^{\boldsymbol{M}}(a,b') = d^{\boldsymbol{M}}(a,b)$ and

$$d^{\mathbb{M}}(a_m \wedge c_i, b'_m \wedge c_i) = |\mathbb{P}^{\mathbb{M}}(a_m \wedge c_i) - \mathbb{P}^{\mathbb{M}}(b'_m \wedge c_i)|,$$

for each $1 \leq i \leq k$. Using the same reasoning as in step 2, "but restricting to c_i ", for each $1 \leq i \leq k$ we can find $(b_{1i}, \dots, b_{ni})$ such that:
(i) $b_{1i} \leq c_i, \dots, b_{ni} \leq c_i$;
(ii) for every m either $a_m \wedge c_i \leq b_{mi}$ or $a_m \wedge c_i > b_{mi}$;
(iii) and $\mathbb{P}^{\mathbb{M}}(b_{1i}) = \mathbb{P}^{\mathbb{M}}(b_1 \wedge c_i), \dots, \mathbb{P}^{\mathbb{M}}(b_{ni}) = \mathbb{P}^{\mathbb{M}}(b_n \wedge c_i)$.
All it remains is to observe that $b' = (b'_1, \dots, b'_n)$ where

$$b'_1 = \bigvee_{1\leq i\leq k} b_{1i}, \dots, b'_n = \bigvee_{1\leq i\leq k} b_{ni}$$

is the required b', which proves the furthermore part.

Step 4. Let C be any small parameter set. We can assume C is a boolean algebra. Let $a \models p, b \models q$ such that $d^{\mathbb{M}}(a,b) = d(p,q)$.

Fix a natural number $k > 0$. By Remark 1.82 there exists a finite set $F \subseteq C$ such that $|d(p,q) - d(p|_F, q|_F)| \leq 1/k$. By Lemma 4.33 by enlarging F we can assume F is a finite boolean algebra such that for each m

$$\|\mathbb{P}^{\mathbb{M}}(a_m|\sigma(C)) - \mathbb{P}^{\mathbb{M}}(a_m|F)\|_1 \leq 1/k,$$
$$\|\mathbb{P}^{\mathbb{M}}(b_m|\sigma(C)) - \mathbb{P}^{\mathbb{M}}(b_m|F)\|_1 \leq 1/k.$$

Therefore

$$\begin{aligned}\max_{1\leq m\leq n} \|\mathbb{P}^{\mathbb{M}}(a_m|\sigma(C)) - \mathbb{P}^{\mathbb{M}}(b_m|\sigma(C))\|_1 &\leq \max_{1\leq m\leq n} \|\mathbb{P}^{\mathbb{M}}(a_m|F) - \mathbb{P}^{\mathbb{M}}(b_m|F)\|_1 + 2/k \\ &\leq d(p|_F, q|_F) + 2/k \leq d(p,q) + 3/k.\end{aligned}$$

Sending $k \to \infty$ we obtain the required inequality. For the furthermore part, by the previous paragraphs, we obtain that

$$r(xy) = p(x) \cup q(y) \cup \{d(x_m, y_m) = \|\mathbb{P}^{\mathbb{M}}(a_m|\sigma(C)) - \mathbb{P}^{\mathbb{M}}(b_m|\sigma(C))\|_1 : 1 \leq m \leq n\},$$

is approximately finitely satisfiable, hence a type. Let $(a, b') \in M_{xy}$ be a realization of r. Then b' is the required element. $\square$

Corollary 4.42. *[BH23, Corollary 6.12] Let $C \subseteq \mathbb{M}$ be a small parameter set and let $a, b \in \mathbb{M}$. Then*

$$d(\mathrm{tp}(a|C), \mathrm{tp}(b|C)) = \|\mathbb{P}^{\mathbb{M}}(a|\sigma(C)) - \mathbb{P}^{\mathbb{M}}(b|\sigma(C))\|.$$

Proof. This follows from applying Theorem 4.41 to the partitions (a, a^c) and (b, b^c) of $1^{\mathbb{M}}$. Note that

$$\begin{aligned}\mathbb{P}^{\mathbb{M}}(a^c|\sigma(C)) - \mathbb{P}^{\mathbb{M}}(b^c|\sigma(C))\|_1 &= \|(1 - \mathbb{P}^{\mathbb{M}}(a|\sigma(C))) - (1 - \mathbb{P}^{\mathbb{M}}(b|\sigma(C)))\|_1 \\ &= \|\mathbb{P}^{\mathbb{M}}(b|\sigma(C)) - \mathbb{P}^{\mathbb{M}}(a|\sigma(C))\|_1\end{aligned}$$

$\square$

Before proving that APA is $\aleph_0$-stable we first establish two lemmas. Recall that $(\Omega, \mathcal{F}, \mathbb{P})$ is an atomless probability space with associated probability algebra $\mathbb{M}$.

Lemma 4.43. *Let $\mathcal{D} \subseteq \mathcal{F}$ be a boolean algebra. Then for every $\varepsilon > 0$ and $A \in \sigma(\mathcal{D})$ there exists $K \in \mathcal{D}$ such that $\mathbb{P}(A \triangle K) \leq \varepsilon$.*

Proof. Let $\mathcal{T}$ be the collection of all $A \in \mathcal{F}$ such that for every $\varepsilon > 0$ there exists $K \in \mathcal{D}$ such that $\mathbb{P}(A \triangle K) \leq \varepsilon$. It is easy to see that $\mathcal{T}$ is a boolean algebra containing $\mathcal{D}$. Fix $\varepsilon > 0$ and let $(A_n) \subseteq \mathcal{T}$ be an increasing sequence of sets, with $A = \bigcup A_n$. Consider m such that $\mathbb{P}(A \triangle A_m) \leq \varepsilon/2$ and $K \in \mathcal{D}$ such that $\mathbb{P}(A_m \triangle K) \leq \varepsilon/2$. We obtain

$$\mathbb{P}(A \triangle K) \leq \mathbb{P}(A \triangle A_m) + \mathbb{P}(A_m \triangle K) \leq \varepsilon.$$

Therefore $\mathcal{T}$ is a σ-algebra, and since it contains $\mathcal{D}$, results that $\sigma(\mathcal{D}) \subseteq \mathcal{T}$. $\square$

Lemma 4.44. *Let $\mathcal{C}$ be a σ-subalgebra of $\mathbb{M}$, which is also a small parameter set. Let $a \in \mathcal{C}$ and $r \in [0,1]$ with $r \leq \mathbb{P}^{\mathbb{M}}(a)$. Then there exists $b \in \mathbb{M}$ with $b \leq a$, such that for every $a' \in \mathcal{C}$ with $a' \leq a$ we have*

$$\mathbb{P}^{\mathbb{M}}(b \wedge a') = r\mathbb{P}^{\mathbb{M}}(a').$$

Proof. Let $p(x) = \{\mathbb{P}(x \wedge a') = r\mathbb{P}(a') : a' \in \mathcal{C}, a' \leq a\}$. Then, using Corollary 4.12, we obtain that $p(x)$ is finitely satisfiable, hence it's a type. Since $\mathbb{M}$ is a monster model, there exists $b \models p$, which is the desired element. $\square$

As an easy corollary, we obtain:

Corollary 4.45. *Let $\mathcal{C}$ be a σ-subalgebra of $\mathbb{M}$, which is also a small parameter set. Let $(a_1, \ldots, a_n) \in \mathcal{C}^n$ be a partition of $1^{\mathbb{M}}$, and $(r_1, \ldots, r_n) \in [0,1]^n$ such that $q_m \leq \mathbb{P}^{\mathbb{M}}(a_m)$ for each $1 \leq m \leq n$. Then there exists $b \in \mathbb{M}$ such that for every $1 \leq m \leq n$ and every $a'_m \leq a_m$ we have $\mathbb{P}^{\mathbb{M}}(b \wedge a'_m) = r_m\mathbb{P}^{\mathbb{M}}(a'_m)$. Moreover, if $\mathcal{D} \subseteq \mathcal{F}$ is a σ-algebra such that $\widehat{\mathcal{D}} = \mathcal{C}$, then*

$$\begin{aligned}\mathbb{P}^{\mathbb{M}}(b|\mathcal{C}) &= r_1\mathbb{P}^{\mathbb{M}}(a_1|\mathcal{C}) + \cdots + r_n\mathbb{P}^{\mathbb{M}}(a_n|\mathcal{C}) \\ &= r_1\chi_{A_1} + \cdots + r_n\chi_{A_n},\end{aligned}$$

where $A_1, \ldots, A_n \in \mathcal{D}$ such that $\widehat{A_1} = a_1, \ldots, \widehat{A_n} = a_1$ and $(A_1, \ldots, A_n)$ forms a partition of Ω.

Proof. Using Lemma 4.44, we find $b_1, \ldots, b_n \in \mathbb{M}$ such that for every $1 \leq m \leq n$ and $a'_m \leq a_m$ we have $\mathbb{P}^{\mathbb{M}}(b_m \wedge a'_m) = r_m\mathbb{P}^{\mathbb{M}}(a'_m)$. Then we can take $b = b_1 \vee \cdots \vee b_n$. Let $A_1, \ldots, A_n \in \mathcal{D}$ such that $\widehat{A_1} = a_1, \ldots, \widehat{A_n} = a_n$ and $(A_1, \ldots, A_n)$ to be a partition of Ω. Then, for every $A \in \mathcal{D}$ we have that

$$\begin{aligned}\int_A \mathbb{P}^{\mathbb{M}}(b|\mathcal{C})\ d\mathbb{P} = \int_A \mathbb{E}(b|\mathcal{D})\ d\mathbb{P} &= \int_A \chi_B\ d\mathbb{P} = \mathbb{P}(B \cap A) \\ &= \mathbb{P}(B \cap (A \cap A_1)) + \ldots \mathbb{P}(B \cap (A \cap A_n)) \\ &= r_1\mathbb{P}(A \cap A_1) + \cdots + r_n\mathbb{P}(A \cap A_n) \\ &= \int_A r_1\chi_{A_1} + \cdots + r_n\chi_{A_n}\ d\mathbb{P} \\ &= \int_A r_1\mathbb{P}^{\mathbb{M}}(a_1|\mathcal{C}) + \cdots + r_n\mathbb{P}^{\mathbb{M}}(a_n|\mathcal{C})\ d\mathbb{P}.\end{aligned}$$

Hence, it follows that $\mathbb{P}^{\mathbb{M}}(b|\mathcal{C}) = r_1\chi_{A_1} + \cdots + r_n\chi_{A_n}$ $\square$

Theorem 4.46. *[BH23, Proposition 6.15] The theory APA is $\aleph_0$-stable.*

Proof. We will apply Proposition 3.19, which says that in order to prove $\aleph_0$-stability of a theory it is enough to prove that the set of 1-types in countably many parameters has a dense set of countable cardinality.

Let $C \subseteq \mathbb{M}$ be a countable parameter set, v a single variable. We need to prove $\mathrm{S}_v(C)$ has a countable dense set in the metric topology. Let $\mathcal{D} \subseteq \mathcal{F}$ be a countable boolean algebra, such that $\widehat{\mathcal{D}} = \mathrm{BA}(C)$, and then $\widehat{\sigma(\mathcal{D})} = \sigma(C)$.

Fix $\varepsilon > 0$ and let $p(x) \in \mathrm{S}_v(C)$. There exists $a \in \mathbb{M}$ with $\mathrm{tp}(a|C) = p(x)$. By Lemma 4.33, there exists $a_1, \dots, a_n \in \sigma(C)$, where $(a_1, \dots, a_n)$ is a partition of $1^{\mathbb{M}}$, such that

$$\|\mathbb{P}^{\mathbb{M}}(a|\sigma(C)) - \mathbb{P}^{\mathbb{M}}(a|\mathrm{BA}(a_1, \dots, a_n))\|_1 \leq \varepsilon.$$

Let $A, A_1, \dots, A_n \in \sigma(\mathcal{D})$ such that $\widehat{A} = a, \widehat{A_1} = a_1, \dots, \widehat{A_n} = a_n$, where $(A_1, \dots, A_n)$ is a partition of Ω. Then

$$\mathbb{P}^{\mathbb{M}}(a|\mathrm{BA}(a_1, \dots, a_n)) = \frac{\mathbb{P}(A \cap A_1)}{\mathbb{P}(A_1)}\chi_{A_1} + \frac{\mathbb{P}(A \cap A_2)}{\mathbb{P}(A_2)}\chi_{A_2} + \cdots + \frac{\mathbb{P}(A \cap A_n)}{\mathbb{P}(A_n)}\chi_{A_n}.$$

By Corollary 4.43 we can find $K_1, \dots, K_n \in \mathcal{D}$ such that $(K_1, K_2, \dots, K_n)$ is a partition of Ω and $\mathbb{P}(A_m \triangle K_m) \leq \varepsilon/n$, for each $1 \leq m \leq n$, and let $q_1, \dots, q_n \in \mathbb{Q}$ such that $0 \leq \frac{\mathbb{P}(A \cap A_m)}{\mathbb{P}(A_m)} - q_m \leq \varepsilon$ for each $1 \leq m \leq n$. By Corollary 4.45 there exists $B \in \mathbb{M}$ such that

$$\mathbb{P}^{\mathbb{M}}(b|\sigma(C)) = \mathbb{E}(B|\sigma(D)) = q_1\chi_{K_1} + \cdots + q_n\chi_{K_n}.$$

Let $\widehat{B} = b$. Then we observe that

$$\begin{aligned}
\|\mathbb{P}^{\mathbb{M}}(a|\sigma(C)) - \mathbb{P}^{\mathbb{M}}(b \mid \sigma(C))\|_1 &\leq \|\mathbb{P}^{\mathbb{M}}(a|\sigma(C)) - \mathbb{P}^{\mathbb{M}}(a|\mathrm{BA}(a_1, \dots, a_n))\|_1 \\
&\quad + \|\mathbb{P}^{\mathbb{M}}(a \mid \mathrm{BA}(a_1, \dots, a_n)) - \mathbb{P}^{\mathbb{M}}(b|\sigma(C))\|_1 \\
&\leq \varepsilon + \sum_{m=1}^{n} \left\| \frac{\mathbb{P}(A \cap A_m)}{\mathbb{P}(A_m)}\chi_{A_m} - q_m\chi_{K_m} \right\|_1 \\
&\leq \varepsilon + \sum_{m=1}^{n} (\varepsilon\mathbb{P}(A_m) + \mathbb{P}(A_m \triangle K_m)) \\
&\leq 3\varepsilon.
\end{aligned}$$

By Corollary 4.42 we have that $d(p, \mathrm{tp}(b|C)) \leq 3\varepsilon$. Note that

$$\mathbb{P}^{\mathbb{M}}(b|\sigma(C)) \in \{q_m\chi_{K_m} + \cdots + q_m\chi_{K_m} : m, q_1, \dots, q_m \in \mathbb{Q}, K_1, \dots, K_m \in \mathcal{D}\} =: S.$$

The set S is obviously countable, from which follows that

$$\mathcal{S} = \{\mathrm{tp}(b|C) : b \in \mathbb{M}, \mathbb{P}^{\mathbb{M}}(b|\sigma(C)) \in S\}$$

is countable, here we also apply Lemma 4.39. Finally, since ε is arbitrary, we obtain that $\mathcal{S}$ is a countable dense subset in the metric topology of $\mathrm{S}_v(C)$. $\square$

4.4 Applications

In this section, we present three applications of the model theory of probability algebras to certain results in probability theory.

We begin with Proposition 4.47, proved in [Hru12, Proposition 2.5], for which Terence Tao later provided a short and elementary proof on his blog [Tao13] using spectral theory. This result has applications in Hrushovski's work on approximate groups, and is used in giving an alternative proof to Tao's algebraic regularity lemma [Tao13].

Proposition 4.47. *For any $p, q \in [0,1]$ with $p < q$ there exists n, such that for any probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and events $A_1, \ldots, A_m, B_1, \ldots, B_m \in \mathcal{F}$ satisfying*

$$\mathbb{P}(A_i \cap B_j) \leq p \text{ and } \mathbb{P}(A_j \cap B_i) \geq q$$

for all $1 \leq i < j \leq m$, then $m \leq n$.

Proof. Suppose the contradiction. Then, for every n there exists a probability algebra $\boldsymbol{M}_n$ and $a_1, \ldots, a_n, b_1, \ldots, b_n \in M_n$ such that $\mathbb{P}^{\boldsymbol{M}_n}(a_i \wedge b_j) \leq p$ and $\mathbb{P}^{\boldsymbol{M}_n}(a_j \wedge b_i) \geq q$ for all $1 \leq i < j \leq n$.

By applying the compactness theorem, it follows that there exists a probability algebra $\boldsymbol{M}$ and sequences $(a_n), (b_n) \subseteq M$ satisfying $\mathbb{P}^{\boldsymbol{M}}(a_i \wedge b_j) \leq p$ and $\mathbb{P}^{\boldsymbol{M}}(a_j \wedge b_i) \geq q$ for all $i < j$. In particular, we obtain $|\mathbb{P}^{\boldsymbol{M}}(a_i \wedge b_j) - \mathbb{P}^{\boldsymbol{M}}(a_j \wedge b_i)| \geq q - p > 0$ for all $i < j$, and as a consequence the formula $\mathbb{P}(x \wedge y)$ (with x, y as single variables) is not stable in $\boldsymbol{M}$. Moreover, by Theorem 4.36, every probability algebra embeds into an atomless probability algebra, hence we can assume $\boldsymbol{M} \models APA$. But, since APA is $\aleph_0$-stable, we obtain that the formula $\mathbb{P}(x \wedge y)$ is stable in every model of APA, contradiction. $\square$

Maharam's Theorem

In this part, we present a proof of Maharam's Theorem which is a fundamental structure theorem about probability algebras. We follow the treatment as in [BH23, Chapter 7]. Let $\boldsymbol{M} \models Pr$.

Definition 4.48. For an atomless element $a \in M$ we define $M{\restriction}a = \{b \in M : b \leq a\}$. Note that $(M{\restriction}a, d^{\boldsymbol{M}})$ is a metric space, where $d^{\boldsymbol{M}}$ is naturally restricted to pairs of elements from $M{\restriction}a$. We define the *density of* a to be the density character of the metric space $(M{\restriction}a, d^{\boldsymbol{M}})$. We call *$a$ homogeneous* if for all $0 < b \leq a$ the density of b equals the density of a. We say a is *maximally homogeneous* if there is no atomless element $b > a$ such that the density of b equals the density of a. We define $\mathcal{K}^{\boldsymbol{M}}$ to be the set of infinite cardinals such that for each $\kappa \in \mathcal{K}$ there is an homogenous element in M of density κ.

Remark 4.49. *The density of an atomless element is an infinite cardinal. If $1_{al}^{\boldsymbol{M}} = 0$ then $\mathcal{K}^{\boldsymbol{M}} = \emptyset$. If instead $0 < 1_{al}^{\boldsymbol{M}}$, it is straightforward to prove that $\mathcal{K}^{\boldsymbol{M}}$ is countable, and that the atomless part can be partitioned as $1_{al}^{\boldsymbol{M}} = \bigvee_{\kappa \in \mathcal{K}^{\boldsymbol{M}}} b_\kappa$ where b_κ is maximally homogeneous of density κ. We then define a function $\Psi^{\boldsymbol{M}} : \mathcal{K}^{\boldsymbol{M}} \to (0,1], \Psi(\kappa) = \mathbb{P}^{\boldsymbol{M}}(b_\kappa)$. If $\mathcal{K}^{\boldsymbol{M}} = \emptyset$, we set $\Psi^{\boldsymbol{M}} = \emptyset$. The elements of the pair $(\mathcal{K}^{\boldsymbol{M}}, \Psi^{\boldsymbol{M}})$ are called the Maharam invariants of $\boldsymbol{M}$.*

The structure of the atomic part $1_{at}^{\boldsymbol{M}}$ of $\boldsymbol{M}$ is uniquely determined by $\mathrm{at}^{\boldsymbol{M}} = (\mathrm{at}^{\boldsymbol{M}}(1))_{n \geq 1}$ (see Proposition 4.20). Consequently, to complete the structure theorem for probability algebras, we need only study the atomless probability algebras.

Definition 4.50. We define $\boldsymbol{M}$ to be *homogeneous* if $1^{\boldsymbol{M}}$ is homogeneous.

Remark 4.51. *In [BH23, Example 7.7] it is proven that if $\boldsymbol{N}$ is the unique separable model of APA, then for any infinite cardinal number κ, the product probability space $\boldsymbol{N}^{\kappa}$ (see [Fre01, Section 254]) is homogeneous and has density κ. Using this result it is easy to prove that for any pair $(\mathcal{K}, \Psi)$, such that $\mathcal{K}$ is a countable set of infinite cardinals and $\Psi : \mathcal{K} \to (0, 1]$ satisfies $\sum_{\kappa \in \mathcal{K}} \Psi(\kappa) = 1$, there exists $\boldsymbol{R} \models APA$ such that its Maharam invariants are exactly $(\mathcal{K}, \Psi)$.*

Before stating the main theorem we need one more definition:

Definition 4.52. Let $\boldsymbol{N}$ be a model of Pr and suppose $\boldsymbol{N}$ is a substructure of $\boldsymbol{M}$. We say that a nonzero element $a \in M$ is an *atom relative to* N if $M{\restriction}a = \{a \wedge b : b \in N\}$, otherwise we say that a is *atomless relative to* N. We say that $\boldsymbol{M}$ is *atomless over* $\boldsymbol{N}$ if every nonzero $a \in M$ is atomless relative to N.

Theorem 4.53 (Maharam's Theorem)**.** *The probability algebra $\boldsymbol{M}$ is uniquely determined, up to isomorphism, by $\mathrm{at}^{\boldsymbol{M}}$ for the atomic part and its Maharam invariants $(\mathcal{K}^{\boldsymbol{M}}, \Psi^{\boldsymbol{M}})$ for the atomless part.*

Proof. It suffices to prove that any two homogeneous models of APA with the same density are isomorphic. The proof will follow from a number of lemmas.

Lemma 1. Let $\boldsymbol{A}$ be a model of Pr, let $B \subseteq A$ be a σ-subalgebra, and let $p(x) \in \mathrm{S}_x(B)$ be a type in the single variable x. Then $\lambda : B \to [0, 1], b \mapsto \mathbb{P}(x \wedge b)^p$ is a finitely additive functional with $\lambda(b) \leq \mathbb{P}^{\boldsymbol{A}}(b)$ for all $b \in B$. Furthermore, $\boldsymbol{A}$ realizes p if and only if there exists an element $a \in A$ such that $\lambda(b) = \mathbb{P}^{\boldsymbol{A}}(a \wedge b)$ for all $b \in B$.

Proof of Lemma 1. The fact that λ is finitely additive can be seen by taking a realization of p in an elementary extension of A. It is obvious that $\lambda \leq \mathbb{P}^{\boldsymbol{A}}$ on the common domain B. The furthermore part follows directly from Lemma 4.39. $\square$

Lemma 2. Let $\boldsymbol{B}, \boldsymbol{A}$ be models of Pr with $\boldsymbol{B} \subseteq \boldsymbol{A}$, and assume that $\boldsymbol{A}$ is atomless over $\boldsymbol{B}$. Then for every finitely additive functional $\lambda : B \to [0, 1]$, such that $\lambda(b) \leq \mathbb{P}^{\boldsymbol{A}}(b)$ for every $b \in B$, there exists an element $a \in \boldsymbol{A}$ such that $\lambda(b) = \mathbb{P}^{\boldsymbol{A}}(a \wedge b)$ for every $b \in B$.

Proof of Lemma 2. Here we present, for completion, the same proof as in [Fre11, Lemma 331B], adapted to our case. First, we observe that λ is a σ-additive functional, as for any increasing chain of elements $(b_n) \subseteq B$ we have

$$\lim_{n\to\infty} \lambda\left(\bigvee_m b_m \setminus \bigvee_{k=0}^{n} b_k\right) \leq \lim_{n\to\infty} \mathbb{P}^{\boldsymbol{A}}\left(\bigvee_m b_m \setminus \bigvee_{k=0}^{n} b_k\right) = 0.$$

For $a \in A$ we define $v_a : B \to [0, 1], v_a(b) = \mathbb{P}^{\boldsymbol{A}}(a \wedge b)$. We prove that for nonzero a there exists a nonzero $c \leq a$ such that $v_c \leq v_a/2$. By assumption, there exists $e \leq a$ such that $e \neq a \wedge b$ for any $b \in B$. Then, by the Hahn Decomposition Theorem for bounded σ-additive functionals [Fre11, 326M], there exists an element $b_0 \in B$ which is a positive set for $v_a - 2v_e$, and b_0^c is a negative set for $v_a - 2v_e$; that is $v_a - 2v_e$ is positive on any smaller element than b_0 and negative on any smaller than b_0^c. If $b_0 \wedge e \neq 0$, then we can consider $c = b_0 \wedge e$ as for all $b \in B$

$$v_c(b) = \mathbb{P}^{\boldsymbol{A}}(b \wedge b_0 \wedge e) = v_e(b \wedge b_0) \leq v_a(b \wedge b_0)/2 \leq v_a(b)/2,$$

while if $b_0 \wedge e = 0$ then we can choose $c = a \backslash (b_0 \vee e)$, which is nonzero as otherwise $e = a \wedge b_0^c$, and we verify that for all $b \in B$

$$v_c(b) = \mathbb{P}^{\boldsymbol{A}}((a \backslash e) \wedge (b \backslash b_0)) = v_a(b \backslash b_0) - v_e(b \backslash b_0) \leq v_a(b \backslash b_0)/2 \leq v_a(b)/2.$$

Then, by induction for any nonzero n exists a nonzero $c \leq a$ such that $v_c \leq v_a/n$.

Let $S = \{a \in A : v_a \leq \lambda\}$. It is a standard fact that any chain in A (with respect to the relation $\leq$) has a supremum in A, and this supremum can be expressed as a countable union of elements from the chain. Notice that S contains $0^{\boldsymbol{A}}$ and is closed under taking suprema of chains. Hence, by Zorn's Lemma we can pick a maximal element $a \in S$.

Suppose that $v_a \neq \lambda$. Since for all $b \in B$, $v_a(b) \leq \lambda(b) \leq \mathbb{P}(b)$, it follows that $\lambda^* = \lambda - v_a \leq v_{a^c}$, and, moreover, $0 < \lambda^*(1) < v_{a^c}(1)$. Therefore, there exists $n > 0$ such that $\lambda^*(1) - v_{a^c}(1)/n > 0$. Applying the Hahn Decomposition Theorem for $\lambda^* - v_{a^c}/n$ yields a nonzero element $b_0 \in B$ with $b_0 \wedge a^c \neq 0$ such that b_0 is a positive set for $\lambda^* - v_{a^c}/n$. Then, by the result established in the paragraph above, there exists a nonzero $c \leq b_0 \wedge a^c$ such that $v_c \leq v_{b_0 \wedge a^c}/n$. It follows that for all $b \in B$

$$v_c(b) \leq v_{b_0 \wedge a^c}(b)/n = v_{a^c}(b \wedge b_0)/n \leq \lambda^*(b \wedge b_0) \leq \lambda^*(b).$$

Since $c \leq a^c$, we have $v_{a \vee c} = v_a + v_c \leq \lambda$. This implies that $a + c \in S$, contradicting the maximality of a. Hence, $v_a = \lambda$. □

Lemma 3. Let $\boldsymbol{A} \models APA$ be homogeneous of density κ. Then A is k-saturated.

Proof of Lemma 3. If k is countable, then $\boldsymbol{A}$ is the unique separable model of APA, which we know to be $\aleph_0$-saturated (see Remark 4.22). If k is uncountable, let $B \subseteq A$ be a parameter set with $|B| < \kappa$. By Lemma 4.43, the Boolean algebra $\mathrm{BA}(B)$ has cardinality $< \kappa$ and is dense in $\sigma(B)$.

For every nonzero $a \in A$, the density character of $A{\restriction}a$ is κ, while the density of the set $\{a \wedge b : b \in \sigma(B)\}$ is at most $|\mathrm{BA}(B)| < \kappa$. It follows that $A{\restriction}a \neq \{a \wedge b : b \in \sigma(B)\}$, so a is atomless over $\sigma(B)$. Consequently, $\boldsymbol{A}$ is atomless over $\sigma(B)$. Then, by Lemma 1 and Lemma 2, it follows that $\boldsymbol{A}$ realizes all types over $\sigma(B)$. □

Finally, if $\boldsymbol{A}, \boldsymbol{B} \models APA$ are homogeneous and have the same density κ, then both are κ-saturated. By the standard back-and-forth method, it follows that $\boldsymbol{A}$ and $\boldsymbol{B}$ are isomorphic. The idea of this method is to first construct a function γ which is an elementary map from $\boldsymbol{A}$ to $\boldsymbol{B}$ such that $\mathrm{dom}(\gamma)$ contains a dense subset of A and $\mathrm{im}(\gamma)$ contains a dense subset of B. Then γ uniquely extends to an isomorphism between $\boldsymbol{A}$ and $\boldsymbol{B}$. □

Atomless Random Variables

In this section, we introduce the theory of *Random Variables* (*RV*) and the theory of *Atomless Random Variables* (*ARV*), folowing the presentation in [BY13]. Our aim is to outline the proof of the equivalence between the stability of APA, the stability of ARV, and a classical probability result due to Ryll-Nardzewski [RN57]. We begin by presenting the probability result.

Definition 4.54. A sequence of $[0,1]$-valued random variables $(\xi_n)_n$ is called *spreadable* if for every n and any increasing sequences of indices $i_1 < \cdots < i_n$ and $j_1 < \cdots < j_n$, the random vectors $(\xi_{i_1}, \ldots, \xi_{i_n})$ and $(\xi_{j_1}, \ldots, \xi_{j_n})$ have the same distribution. We also call the sequence $(\xi_n)_n$ *exchangeable* if for every n and any sequences of distinct indices $i_1, \ldots, i_n$ and $j_1, \ldots, j_n$ the random variables $(\xi_{i_1}, \ldots, \xi_{i_n})$ and $(\xi_{j_1}, \ldots, \xi_{j_n})$ have the same distribution.

Theorem 4.55 (Ryll-Nardzewski Theorem [RN57])**.** *A sequence of* $[0,1]$*-valued random variables is spreadable if and only if it is exchangeable.*

We are particularly interested in the atomless variant of the theorem:

Every sequence of $[0,1]$*-valued random variables defined on an atomless probability space is spreadable if and only if it is exchangeable.*

It can be shown that this statement is equivalent to Ryll-Nardzewski theorem. Indeed, for any probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the product space

$$(\Omega \times [0,1],\ \mathcal{F} \otimes \mathrm{Borel}([0,1]),\ \mathbb{P} \times \mu)\,,$$

where μ denotes the Lebesgue measure on $[0,1]$, is an atomless probability space. To see this, note that if for some set $A \in \mathcal{F} \otimes \mathrm{Borel}([0,1])$ and positive integer n, we have $\frac{1}{n} < \mathbb{P} \times \mu(A)$, then by subdividing $[0,1]$ into n equal subintervals, there must exist some $0 \leq m < n$ such that

$$0 < \mathbb{P} \times \mu\left(A \cap \left(\Omega \times \left[\frac{m}{n}, \frac{m+1}{n}\right]\right)\right) \leq \frac{1}{n}.$$

Given any sequence of $[0,1]$-valued random variables $(\xi_n)_n$ on $(\Omega, \mathcal{F}, \mathbb{P})$, we define a corresponding sequence of $[0,1]$-valued random variables $(\xi'_n)_n$ on the product space by

$$\xi'_n : \Omega \times [0,1] \to [0,1],\ \xi'_n(\omega, r) := \xi_n(\omega).$$

This construction preserves spreadability and exchangeability, so that (ξ_n) is spreadable (respectively, exchangeable) if and only if (ξ'_n) is spreadable (respectively, exchangeable). Hence, Ryll-Nardzewski Theorem follows directly from its atomless variant.

Before defining the theory RV and showing the equivalence stated in the introduction we need a couple of preliminary propositions. Let $\mathcal{L}, \mathcal{G}$ be two continuous languages, $\boldsymbol{M}$ be an $\mathcal{L}$-structure, $\boldsymbol{N}$ be a $\mathcal{G}$-structure, T be an $\mathcal{L}$-theory, and x, y be multivariables.

Definition 4.56. For a total order I, a sequence of elements $(a_i)_{i\in I} \subseteq M_x$ is called *indiscernible* if for every n and any increasing sequences of indices $i_1 < \cdots < i_n$ and $j_1 < \ldots j_n$

$$\operatorname{tp}(a_{i_1},\ldots,a_{i_n}) = \operatorname{tp}(a_{j_1},\ldots,a_{j_n}).$$

We also call $(a_i)_{i\in I}$ *totally indiscernible* if for every n and any sequences of distinct indices $i_1,\ldots,i_n$ and $j_1,\ldots,j_n$

$$\operatorname{tp}(a_{i_1},\ldots,a_{i_n}) = \operatorname{tp}(a_{j_1},\ldots,a_{j_n}).$$

The following is a classical result of model theory which also holds in the continuous case:

Theorem 4.57. *[Pil83, Proposition 7.1] The theory T is stable if and only if in every model of T every indiscernible sequence indexed by $\mathbb{N}$ is totally indiscernible.*

Proof. Suppose T is unstable. Then there exists an $\mathcal{L}(x,y)$-formula φ, an $\mathcal{L}$-structure $\boldsymbol{M} \models T$, a real number $\varepsilon > 0$ and sequences $(a_n) \subseteq M_x, (b_n) \subseteq M_y$ such that $|\varphi^{\boldsymbol{M}}(a_m,b_n) - \varphi^{\boldsymbol{M}}(a_n,b_m)| > \varepsilon$ for all $m<n$. Note that, by taking $\psi(xyx'y') := \varphi(yx')$, for x',y' multivariables of the same sort as x,y, and defining the sequence $(c_n)_n = (a_nb_n)_n$, then we have that $|\psi^{\boldsymbol{M}}(c_m,c_n) - \psi^{\boldsymbol{M}}(c_n,c_m)| > \varepsilon$ for every $m<n$. By applying $\aleph_0$-infinite Ramsey's Theorem we can assume, without loss of generality that $\psi^{\boldsymbol{M}}(c_m,c_n) \leq r < s \leq \psi^{\boldsymbol{M}}(c_n,c_m)$ for all $m<n$.

Now, for a sequence of multivariables (x_n) of sort xy, let Σ be the following set of conditions

(i) $\psi(x_m,x_n) \leq r$ and $\psi(x_n,x_m) \geq s$ for all $m<n$,

(ii) $|\gamma(x_{i_1},\ldots,x_{i_n}) - \gamma(x_{j_1},\ldots,x_{j_n})| = 0$ for all $\mathcal{L}$-formulas γ and sequences of indices $i_1 < \cdots < i_n$ and $j_1 < \cdots < j_n$.

Using $\aleph_0$-infinite Ramsey's Theorem it is straightforward to see that Σ is finitely approximately satisfiable in $\boldsymbol{M}$. Let $(d_n)_n$ be a realization of Σ in some elementary extension $\boldsymbol{N} \succcurlyeq \boldsymbol{M}$. Then, by construction, (d_n) is an indiscernible sequence and

$$\psi^{\boldsymbol{N}}(d_m,d_n) \leq r,\ s \leq \psi^{\boldsymbol{N}}(d_n,d_m)$$

for all $m<n$. Therefore $(d_n)_n$ is not totally indiscernible.

For the converse, suppose a sequence (b_n) of a model $\boldsymbol{M} \models T$ is indiscernible but not totally indiscernible. Then, by compactness, there exists an indiscernible sequence $(a_i)_{i\in\mathbb{Q}}$ which is not totally indiscernible, in some elementary extension of $\boldsymbol{M}$, which, by abuse of notation, we continue to denote by $\boldsymbol{M}$. Then, since permutations are generated by traspositions, and up to reindexing, we can assume there exists an $\mathcal{L}$-formula γ such that

$$\gamma^{\boldsymbol{M}}(a_0,\ldots,a_{k-1},a_{k+2}\ldots a_n,a_k,a_{k+1}) = r,\ \gamma^{\boldsymbol{M}}(a_0,\ldots,a_{k-1},a_{k+2}\ldots a_n,a_{k+1},a_k) = s$$

with $r \neq s$. Then, if we denote $c = a_0\ldots a_{k-1}a_{k+2}\ldots a_n$, let $(d_n)_n = (ca_{k+1-1/2^n})_n$ and $\varphi(xyxy') = \gamma(xyy')$ with x of sort c, then $|\varphi^{\boldsymbol{M}}(d_m,d_n) - \varphi^{\boldsymbol{M}}(d_n,d_m)| = |r-s| > 0$ for every $m<n$. Therefore φ is not stable, and as a consequence T is not stable. $\square$

Definition 4.58. We call a tuple $(\mathcal{L}', \boldsymbol{M}', D, \iota, (f')_{f\in\mathcal{G}^f}, (R')_{R\in\mathcal{G}^r}, d')$ *an intepretation of* $\boldsymbol{N}$ in $\boldsymbol{M}$ if $\mathcal{L}'$ is a language extension of $\mathcal{L}$, $\boldsymbol{M}'$ is an $\mathcal{L}'$-structure that is an expansion by definitions of $\boldsymbol{M}$ such that

(i) D is a 0-definable set in $\boldsymbol{M}'$ and ι is a bijection $N \to D$,

(ii) for every n-ary function symbol $f \in \mathcal{G}^f$, f' is a function symbol of $\mathcal{L}'$ satisfying

$$(f')^{\boldsymbol{M}'}(\iota(a)) = \iota(f^{\boldsymbol{N}}(a)) \text{ for all } a \in N^n,$$

(iii) for every n-ary relation symbol $R \in \mathcal{G}^r$, R' is a relation symbol of $\mathcal{L}'$ satisfying

$$(R')^{\boldsymbol{M}'}(\iota(a)) = R^{\boldsymbol{M}}(a) \text{ for all } a \in N^n,$$

(iv) d' is a binary relation symbol of $\mathcal{L}'$ satisfying

$$(d')^{\boldsymbol{M}'}(\iota(a), \iota(b)) = d^{\boldsymbol{N}}(a, b) \text{ for all } a, b \in N.$$

We also say that $f^{\boldsymbol{N}}, R^{\boldsymbol{N}}, d^{\boldsymbol{N}}$ are *defined by* or *represented by* f', R', d', respectively. Finally, we call $\boldsymbol{N}$ *interpretable* in $\boldsymbol{M}$ if there exists an interpretation of $\boldsymbol{N}$ in $\boldsymbol{M}$. In practice, we will not explicitly construct the expanded structure by definitions; instead, we will specify the $\mathcal{L}$-formulas that define the new symbols of $\mathcal{L}'$ and say that these formulas define or represent $f^{\boldsymbol{N}}, R^{\boldsymbol{N}}$ and $d^{\boldsymbol{N}}$.

Proposition 4.59. *Suppose* $\boldsymbol{N}$ *is interpretable in* $\boldsymbol{M}$ *by the intepretation*

$$(\mathcal{L}', \boldsymbol{M}', D, \iota, (f')_{f\in\mathcal{G}^f}, (R')_{R\in\mathcal{G}^r}, d').$$

Then for every $\mathcal{G}(x)$*-formula* φ *there exists an* $\mathcal{L}'(x)$*-formula* φ' *such that* $\varphi^{\boldsymbol{N}}(a) = (\varphi')^{\boldsymbol{M}'}(\iota(a))$ *for all* $a \in N_x$*.*

Proof. The formula φ' is obtained by replacing each occurence of $f, R, d, \inf, \sup$ in φ with $f', R', d', \inf_D, \sup_D$ (we can quantify over definable sets, see Remark 2.34). $\square$

Note that in Definition 3.5, if we replace φ with a 0-definable predicate P of $\boldsymbol{M}$, then we obtain the definition for P to be *ε-stable in* $\boldsymbol{M}$ and for P to be *stable in* $\boldsymbol{M}$. It is easy to see that if every $\mathcal{L}$-formula is stable in $\boldsymbol{M}$ then every 0-definable predicate is stable in $\boldsymbol{M}$.

Definition 4.60. We call $\boldsymbol{M}$ *stable* if every $\mathcal{L}$-formula φ is stable in $\boldsymbol{M}$.

Proposition 4.61. *Let* $\boldsymbol{M}'$ *be an* $\mathcal{L}'$*-structure with* $\mathcal{L} \subseteq \mathcal{L}'$*, obtained as an expansion by definitions of* $\boldsymbol{M}$*. If* $\boldsymbol{M}$ *is stable, then* $\boldsymbol{M}'$ *is also stable.*

Proof. Let φ be an $\mathcal{L}'$-formula. By Theorem 2.49, there exists a 0-definable predicate P in $\boldsymbol{M}$ such that $\boldsymbol{M}' \models P = \varphi$. Since P is stable in $\boldsymbol{M}$, it follows that φ is stable in $\boldsymbol{M}'$. $\square$

Corollary 4.62. *Suppose $\boldsymbol{N}$ is interpretable in $\boldsymbol{M}$. If $\boldsymbol{M}$ is stable then $\boldsymbol{N}$ is stable.*

Proof. Let $\boldsymbol{M}'$ be the expansion by definitions of $\boldsymbol{M}$ for which there exists an interpretation of $\boldsymbol{N}$. By Proposition 4.61, $\boldsymbol{M}'$ is stable. Suppose some $\mathcal{G}$-formula φ is unstable in $\boldsymbol{N}$. Then, by Proposition 4.59, it follows that φ' is unstable in $\boldsymbol{M}'$, contradiction. Therefore $\boldsymbol{N}$ is stable. □

We conclude by introducing one final notion, that of *imaginaries*. A rigorous definition and a detailed analysis of its properties are not possible within our current framework, since this would require developing continuous logic for languages with more than one sort, while our treatment has been restricted to single-sorted languages. Nonetheless, we offer an informal outline of its construction and principal properties. For a precise and rigorous account, the reader is referred to Section 5 of [BYU10].

In continuous logic the construction for imaginaries mirrors the same construction as in classical model theory. Let x be a finite multivariable and y be a possible infinite multivariable. Let P be an 0-definable xy-predicate in $\boldsymbol{M}$. Then for $a, b \in M_y$ we set $a \sim b$ if $P(\cdot, a) = P(\cdot, b)$, these are functions $M_x \to [0,1]$. Then, to the language $\mathcal{L}$ we add one more sort for imaginaries, and to $\boldsymbol{M}$ we add the set $M_P = M_y/\sim$ (for the imaginary sort), called the set of canonical parameters for $P(x, y)$, and call this structure $\boldsymbol{M}_P$. The distance on M_P is $d_P(\tilde{a}, \tilde{b}) = \sup_{c \in M_x} |P(c, a) - P(c, b)|$ for $a, b \in M_y$, where $\tilde{a}$ represents its class in M_P. To be more precise, for the imaginary sort, the underlying set is actually taken the completion of $M_y/\sim$ with respect to the distance d_P, as all the metrics for each sort must be complete. Lastly, besides the symbol d_P the language L is also extended with a relation symbol $\widetilde{P}$, such that its interpretation is

$$\widetilde{P} : M_x \times M_P \to [0,1], (c, \tilde{a}) \mapsto P(c, a).$$

Therefore, to summarize, by adding an imaginary sort for the predicate P we make the following extensions:

(i) a new sort is added to represent the set $M_P = M_y/\sim$,

(ii) a new distance symbol d_P such that $d_P(\tilde{a}, \tilde{b}) = \sup_{c \in M_x} |P(c, a) - P(c, b)|$ and a new relation symbol $\widetilde{P}$ such that $\widetilde{P} : M_x \times M_P \to [0,1], (c, \tilde{a}) \mapsto P(c, a)$.

The structure $\boldsymbol{M}_P$ is called the *imaginary expansion of $\boldsymbol{M}$ by P*. We will also use the following standard fact:

Remark 4.63. *In continuous model theory, as in classical model theory, if $\boldsymbol{M}$ is stable, then so is $\boldsymbol{M}_P$.*

We define the *language of random variables* by $\mathcal{L}_{RV} = \{0, \neg, \frac{1}{2}, \dot{-}\}$, where 0 is a constant symbol, $\neg$ and $\frac{1}{2}$ are unary function symbols, and $\dot{-}$ is a binary function symbol. For convenience, we allow a slight abuse of notation, as the set $\{0, \neg, \frac{1}{2}, \dot{-}\}$ is also used to denote continuous connectives. The intended interpretation is always clear from context: when any of these symbols occurs between *terms* in a formula, it is to be understood as a symbol of $\mathcal{L}_{RV}$; when it occurs between *formulas*, it is interpreted as a continuous connective. We also use the abbreviations: 1 for $\neg 0$ and $E(x)$ for $d(x,0)$.

The theory of random variables RV is axiomatized by the following axioms:

(i) $E(x) = E(x \dot{-} y) + E(x \wedge y)$,

(ii) $E(1) = 1$,

(iii) $d(x,y) = E(x \dot{-} y) + E(y \dot{-} x)$,

(iv) $(x \dot{-} y) \dot{-} x = 0,\ ((x \dot{-} z) \dot{-} (x \dot{-} y)) \dot{-} (y \dot{-} z) = 0,\ (x \wedge y) \dot{-} (y \wedge x) = 0$,

(v) $(x \dot{-} y) \dot{-} (\neg y \dot{-} \neg x) = 0,\ \frac{1}{2}x \dot{-} (x \dot{-} \frac{1}{2}x) = 0, (x \dot{-} \frac{1}{2}x) \dot{-} \frac{1}{2}x = 0$.

The axioms above are stated succinctly; the intended interpretation is that all free variables are universally quantified. Moreover, for any two formulas ψ, ϕ, the notation $\psi = \phi$ abbreviates the formula $d(\psi, \phi) = 0$. Similarly, for any two terms t, t', the expression $t = t'$ stands for $d(t,t') = 0$. For a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ we denote by $L^1((\Omega, \mathcal{F}, \mathbb{P}), [0,1])$ the closed subspace of $[0,1]$-valued random variables of $L^1(\Omega, \mathcal{F}, \mathbb{P})$. We now state the following important theorem:

Theorem 4.64. *[BY13, Corollary 2.8] An $\mathcal{L}_{RV}$-structure $\boldsymbol{M}$ is a model of RV if and only if $\boldsymbol{M}$ is isomorphic to $L^1((\Omega, \mathcal{F}, \mathbb{P}), [0,1])$ for a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.*

In Theorem 4.64, the interpretation for the $\mathcal{L}_{RV}$-symbols in the $\mathcal{L}_{RV}$-structure $\boldsymbol{M} = L^1((\Omega, \mathcal{F}, \mathbb{P}), [0,1])$ is the natural one, that is, by letting ω vary over Ω, for two random variables X and Y in $L^1((\Omega, \mathcal{F}, \mathbb{P}), [0,1])$:

(i) $0^{\boldsymbol{M}}$ is the zero random variable,

(ii) $\neg^{\boldsymbol{M}} X$ is the variable defined by $\neg^{\boldsymbol{M}} X(\omega) = 1 - X(\omega)$,

(iii) $\frac{1}{2}^{\boldsymbol{M}} X$ is defined by $\frac{1}{2}^{\boldsymbol{M}} X(\omega) = \frac{X(\omega)}{2}$,

(iv) $X \dot{-}^{\boldsymbol{M}} Y$ is the random variable defined by $X \dot{-}^{\boldsymbol{M}} Y(\omega) = \max(0, X(\omega) - Y(\omega))$.

Finally, it is straightforward to verify that these definitions are well defined up to the equivalence relation identifying two random variables whenever they differ only on a set of measure zero.

For a model $\boldsymbol{M} \models RV$ the proof of Theorem 4.64 (in [BY13]) uses truth assignments, which are maps $\lambda : M \to [0,1]$ which respects $\neg, \frac{1}{2}, \dot{-}$. The set of all truth assignments $M \to [0,1]$ is denoted by $\widetilde{M}$. In [BY13], the set $\widetilde{M}$ is identified as a subset of the product space $[0,1]^M$, endowed with the product topology. It is shown that $\widetilde{M}$, equipped with the subspace topology, is compact and hausdorff. Furthermore, it is proven that there exists a Borel probability measure $\mathbb{P}$ on $\widetilde{M}$ such that $\boldsymbol{M}$ is isomorphic to $L^1((\widetilde{M}, \mathcal{F}, \mathbb{P}), [0,1])$, where $\mathcal{F}$ is the Borel σ-algebra of $\widetilde{M}$.

Remark 4.65. *Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\boldsymbol{M} = L^1((\Omega, \mathcal{F}, \mathbb{P}), [0,1])$, notice that for $a \in \Omega$ the map $\lambda_a : M \to [0,1], X \mapsto X(a)$ is a truth assignment. Therefore, in this case, we can view Ω as a subset of $\widetilde{M}$.*

We call

$$\sup_x \inf_y \left(E(y \wedge \neg y) \vee \left| E(x \wedge y) - \frac{E(x)}{2} \right| \right) = 0.$$

the *atomless random variable axiom*, denoted ARV. We call the theory axiomatized by *RV* together with the axiom ARV as the *atomless random variable theory*, denoted by *ARV*.

Proposition 4.66. *[BY13, Corollary 2.8] Let $\boldsymbol{M} \models RV$. Then $\boldsymbol{M} \models ARV$ if and only if $\boldsymbol{M}$ is isomorphic to $L^1((\Omega, \mathcal{F}, \mathbb{P}), [0,1])$ for some atomless probability space $(\Omega, \mathcal{F}, \mathbb{P})$.*

This proposition has an easy proof using Corollary 4.12.

Theorem 4.67. *[BY13, Propositions 2.11, 2.12] Every model of APA is interpretable in a model of ARV, and every model of ARV is interpretable in an imaginary expansion of a model of APA.*

Proof. Let $\boldsymbol{M} \models APA$ and consider $\boldsymbol{M}$ to be the probability algebra associated with an atomless probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\boldsymbol{N} = L^1((\Omega, \mathcal{F}, \mathbb{P}), [0,1])$, which is a model of ARV. Let $D = \{\chi_a : a \in M\} \subseteq N$, where $\chi_a = \chi_A$ for $\widehat{A} = a$ with $A \in \mathcal{F}$. Hence, it is clear that M is in bijection with D, via the map that sends $a \in M$ to $\chi_a \in D$. Moreover, D is definable in $\boldsymbol{N}$ because for any $f \in N$ we have

$$d^{\boldsymbol{N}}(f, D) = E(f \wedge \neg f),$$

as $d^{\boldsymbol{N}}(f, \chi_A) = \int |f - \chi_A| d\mathbb{P}$ is minimized for $A = f^{-1}(1/2, 1]$. For the function symbols, $(\cdot^c)^{\boldsymbol{M}}$ is defined by $\neg(x)$, $\vee^{\boldsymbol{M}}$ is defined by $x \dotplus (y \dot- x)$. Finally, $d^{\boldsymbol{M}}$ is represented by $d(x,y)$, and $\mathbb{P}^{\boldsymbol{M}}$ is represented by $E(x)$.

Now, let $\boldsymbol{N} = L^1((\Omega, \mathcal{F}, \mathbb{P}), [0,1])$ for an atomless probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $\boldsymbol{M}$ as its probability algebra. Let $D = \{\frac{m}{2^n} : 0 < n, 0 < m < 2^n\}$ and $D' = D \cup \{0,1\}$. Then, given a sequence of variables $X = (x_r)_{r \in D}$, we define by recursion on the smallest $m > 0$ such that $mr \in \mathbb{N}$, a family of terms by $\tau_r = (x_r \vee \tau_{r+1/2^m}) \wedge \tau_{r-1/2^m}$, starting from $\tau_0 = 1$ and $\tau_1 = 0$. Let $\tau(X) = (\tau_r(X))_{r \in D}$. Observe that for $f \in N$ we have

$$\tau((\widehat{f^{-1}[r,1]})_{r \in D}) = (\widehat{f^{-1}[r,1]})_{r \in D}.$$

Then, we define

$$\varphi_n(y, X) = \sum_{k \leq 2^n} 2^{-n} \mathbb{P}(y \wedge \tau_{k/2^n}).$$

It is straightforward to show that the sequence of formulas $(\varphi_n)_n$ is Cauchy. Let P denote its limit. Then, for any $a \in M$, we have

$$P\left(a, \left(\widehat{f^{-1}[r,1]}\right)_{r \in D}\right) = \int \chi_a f \, d\mathbb{P}.$$

Finally, we consider the imaginary expansion $\boldsymbol{M}_P$. Define a map $\iota : N \to M_P$ by

$$\iota(f) = \left(\widehat{f^{-1}[r,1]}\right)_{r \in D}.$$

This map is a bijection and is well-defined up to identification of random variables that agree almost everywhere.

We first show that the function $\dot{-}^{\boldsymbol{N}}$ is represented in $\boldsymbol{M}_P$ by a formula $\gamma(x, y, z)$, where x, y, z are variables of sort M_P. Specifically, γ defines a function on M_P such that for any $f, g, h \in N$,

$$\gamma^{\boldsymbol{M}_P}(\iota(f), \iota(g), \iota(h)) = d_P(\iota(f \dot{-} g), \iota(h)).$$

Observe that

$$\begin{aligned}
d_P(\iota(f \dot{-} g), \iota(h)) &= \sup_a \left| \widetilde{P}(a, \iota(f \dot{-} g)) - \widetilde{P}(a, \iota(h)) \right| \\
&= \sup_a \left| \int \chi_a (f \dot{-} g)\, d\mathbb{P} - \int \chi_a h\, d\mathbb{P} \right| \\
&= \sup_a \left| \sup_b \left(\int \chi_{a \wedge b} f\, d\mathbb{P} \dot{-} \int \chi_{a \wedge b} g\, d\mathbb{P} \right) - \int \chi_a h\, d\mathbb{P} \right| \\
&= \sup_a \left| \sup_b \left(\widetilde{P}(a \wedge b, \iota(f)) \dot{-} \widetilde{P}(a \wedge b, \iota(g)) \right) - \widetilde{P}(a, \iota(h)) \right|.
\end{aligned}$$

Motivated by this, for variables v, w of sort M, we define the formula

$$\gamma(x, y, z) = \sup_v \left| \sup_w \left(\widetilde{P}(v \wedge w, x) \dot{-} \widetilde{P}(v \wedge w, y) \right) - \widetilde{P}(v, z) \right|,$$

which represents $\dot{-}^{\boldsymbol{N}}$. Similarly

(i) the constant $0^{\boldsymbol{N}}$ is defined by $\gamma_1(x) = \widetilde{P}(1, x)$,

(ii) the function $\neg^{\boldsymbol{N}}$ is defined by $\gamma_2(x, y) = \sup_v |(\mathbb{P}(v) \dot{-} \widetilde{P}(v, x)) - \widetilde{P}(v, y)|$,

(iii) the function $\frac{1}{2}^{\boldsymbol{N}}$ is defined by $\gamma_3(x, y) = \sup_v |\widetilde{P}(v, x)/2 - \widetilde{P}(v, y)|$,

(iv) and the distance $d^{\boldsymbol{N}}$ is defined by

$$\gamma_4(x, y) = \sup_v (\widetilde{P}(v, x) \dot{-} \widetilde{P}(v, y)) \dot{+} \sup_v (\widetilde{P}(v, y) \dot{-} \widetilde{P}(v, x)).$$

For example, for every $f, g \in N$ we have

$$d^{\boldsymbol{N}}(f, g) = \int |f - g| d\mathbb{P} = \int \max(0, f - g) d\mathbb{P} + \int \max(0, g - f) d\mathbb{P} = \gamma_4^{\boldsymbol{M}_P}(\iota(f), \iota(g)).$$

This completes the intepretation of $\boldsymbol{N}$ in $\boldsymbol{M}_P$. $\square$

Remark 4.68. *The atomic formulas of $\mathcal{L}_{RV}$ are of the form $d(t,t')$ for some $\mathcal{L}_{RV}$-terms t,t'. And since $RV \models d(t,t') = E(t \dot{-} t') \dot{+} E(t' \dot{-} t)$, we conclude that a quantifier-free type $p(x)$ over a model of RV only depends on conditions of the type $E(t) = E(t)^p$.*

We also need the following important theorem.

Proposition 4.69. *[BY13, Theorem 2.17] (i) The theory ARV has QE.*
(ii) Let $\boldsymbol{M} \models ARV$, with $\boldsymbol{M} = L^1((\Omega, \mathcal{F}, \mathbb{P}), [0,1])$, and let $(f_1, \ldots, f_n)$ and $(g_1, \ldots, g_n)$ be tuples in M^n. Then,

$$\mathrm{tp}(f_1, \ldots, f_n) = \mathrm{tp}(g_1, \ldots, g_n)$$

if and only if the random vectors $(f_1, \ldots, f_n)$ and $(g_1, \ldots, g_n)$ have the same distribution.

Proof. The proof of $(i), (ii)$ is given in [BY13, Theorem 2.17]. For completeness, we assume (i) and give a prove of (ii). By quantifier elimination, every type $p(x) = p(x_1, \ldots, x_n)$ in ARV is determined by its quantifier-free conditions. In particular, by Remark 4.68, it follows that p is completely determined by the values $E(t)^p$ for all $\mathcal{L}_{RV}(x)$-terms t. Note that, since t is constructed from $\{0, \neg, \frac{1}{2}, \dot{-}\}$, it can be viewed as a continuous function $[0,1]^n \to [0,1]$. Consequently, the type $\mathrm{tp}(f_1, \ldots, f_n)$ is uniquely determined by the values

$$E(t(f_1, \ldots, f_n))^{\boldsymbol{M}} = \int t(f_1, \ldots, f_n)\, d\mathbb{P}.$$

Therefore, if the random vectors $(f_1, \ldots, f_n)$ and $(g_1, \ldots, g_n)$ have the same distribution, then for every $\mathcal{L}_{RV}(x)$-term t

$$\int t(f_1, \ldots, f_n)\, d\mathbb{P} = \int t(g_1, \ldots, g_n)\, d\mathbb{P}. \tag{4.1}$$

It follows that $\mathrm{tp}(f_1, \ldots, f_n) = \mathrm{tp}(g_1, \ldots, g_n)$. Conversely, suppose $\mathrm{tp}(f_1, \ldots, f_n) = \mathrm{tp}(g_1, \ldots, g_n)$. Then equality (4.1) holds for every $\mathcal{L}_{RV}(x)$-term t. Since, by Proposition 1.4, the set of functions t is dense in $\mathrm{C}([0,1]^n, [0,1])$, it is an easy exercise to prove that $(f_1, \ldots, f_n)$ and $(g_1, \ldots, g_n)$ have the same distribution. □

Theorem 4.70. *The following are equivalent:*

(i) APA is stable;

(ii) ARV is stable;

(iii) (Ryll-Nardzewski) Every spreadable sequence of [0,1]-valued random variables is exchangeable.

Proof. The equivalence $(i) \leftrightarrow (ii)$ follows from Theorem 4.67 and Remark 4.63. The equivalence $(ii) \leftrightarrow (iii)$ follows from Proposition 4.69 and Theorem 4.57. □

Finally, since APA is stable by Theorem 4.46, the preceding theorem establishes the stability of ARV, and provides an alternative proof of the Ryll-Nardzewski theorem.

Bibliography

[BBHU08] I. Ben Yaacov, A. Berenstein, C. W. Henson, and A. Usvyatsov, Model theory for metric structures, *London Math. Soc. Lecture Note Ser.*, 350:315, 2008.

[BH23] A. Berenstein and C. W. Henson, Model theory of probability spaces, in *Model Theory of Operator Algebras, Ser. Logic Appl.*, vol. 11, pp.159–213, 2023.

[BY13] I. Ben Yaacov, On theories of random variables, *Israel J. Math.*, 194:957–1012, 2013.

[BY14] I. Ben Yaacov, Model theoretic stability and definability of types, after A. Grothendieck, *Bull. Symbolic Logic*, 20(4):491–496, 2014.

[BYI25] I. Ben Yaacov and T. Ibarlucía, *Stability in affine logic*, arXiv:2503.10506, 2025, `https://arxiv.org/abs/2503.10506`.

[BYU10] I. Ben Yaacov and A. Usvyatsov, Continuous first order logic and local stability, *Trans. Amer. Math. Soc.*, 362(10):5213–5259, 2010.

[CK66] C.-C. Chang and H. J. Keisler, *Continuous Model Theory*, Annals of Mathematics Studies, vol. 58, Princeton Univ. Press, Princeton, NJ, 1966.

[Con19] J. B. Conway, *A course in functional analysis*, vol. 96, Springer, 2019.

[Dud18] R. M. Dudley, *Real analysis and probability*, Chapman and Hall/CRC, 2018.

[Dur19] R. Durrett, *Probability: theory and examples*, vol. 49, Cambridge Univ. Press, 2019.

[Fre01] D. H. Fremlin, *Measure Theory*, vol. 2, Torres Fremlin, 2001.

[Fre11] D. H. Fremlin, *Measure Theory*, vol. 3, Torres Fremlin, 2011.

[GK20] I. Goldbring and H. J. Keisler, Continuous sentences preserved under reduced products, *J. Symbolic Logic*, 87(2):649–681, 2020.

[Gro52] A. Grothendieck, Critères de compacité dans les espaces fonctionnels généraux, *Amer. J. Math.*, 74(1):168–186, 1952.

[Hal09] P. Halmos and S. Givant, *Introduction to Boolean algebras*, Springer, 2009.

[Hru12] E. Hrushovski, Stable group theory and approximate subgroups, *J. Amer. Math. Soc.*, 25(1):189–243, 2012.

[Pil83] A. Pillay, *An Introduction to Stability Theory*, Oxford Logic Guides, vol. 8, Oxford Univ. Press, 1983.

[Pil96] A. Pillay, *Geometric stability theory*, Oxford Univ. Press, 1996.

[Roy10] H. L. Royden and P. M. Fitzpatrick, *Real analysis*, 4th ed., Prentice-Hall, Boston, 2010.

[RN57] C. Ryll-Nardzewski, On stationary sequences of random variables and the de Finetti's equivalence, *Colloq. Math.*, 4(2):149–156, 1957.

[SS95] L. A. Steen and J. A. Seebach Jr., Counterexamples in topology, Dover Publications, Mineola, 1995.

[Tao13] T. Tao, A spectral theory proof of the algebraic regularity lemma, `https://terrytao.wordpress.com/2013/10/29/a-spectral-theory-proof-of-the-algebraic-regularity-lemma/`, 2013.

[WH07] C. Ward Henson and H. Tellez, Algebraic closure in continuous logic, *Rev. Colombiana Matem.*, 41:279–285, 2007.

[Wil12] S. Willard, *General topology*, Courier Corporation, 2012.