Continuous Logic and Probability Algebras

THESIS

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in the Graduate School of The Ohio State University

By
Fan Yang, B.S.

Graduate Program in Mathematics

The Ohio State University
2016

Master's Examination Committee:

Christopher Miller, Advisor
Timothy Carlson
Copyright by

Fan Yang

2016
Continuous logic is a multi-valued logic where the set of truth values is the unit interval \([0, 1]\). It was developed recently as a framework for metric structures, which consist of complete, bounded metric spaces on which there are distinguished elements and uniformly continuous functions. There are many parallels between continuous logic and first-order logic: 0 corresponds to true and 1 to false, sup and inf take the place of quantifiers, and uniformly continuous functions on \([0, 1]\) replace connectives. Instead of a distinguished equality symbol, there is a distinguished predicate for the metric. Familiar theorems of first-order logic, such as completeness, compactness, and downward Löwenheim-Skolem, have modified counterparts in continuous logic. We present these results in comparison to those in first order logic and prove that the class of probability algebras (probability spaces modulo null sets where the distance between two events is their symmetric difference) is axiomatizable in continuous logic.
Vita

2009 ...................................................... Bozeman High School

2013 ...................................................... B.S. Mathematics, Carnegie Mellon University

2013--2015 ............................................. Graduate Teaching Associate, Department of Mathematics, The Ohio State University

Fields of Study

Major Field: Mathematics
# Table of Contents

Abstract ................................................................................................................................. ii

Vita ........................................................................................................................................ iii

Chapter 1: Introduction ................................................................................................. 1

Chapter 2: Preliminaries ............................................................................................... 3

Chapter 3: Axioms of continuous logic ........................................................................ 9

Chapter 4: Probability algebras .................................................................................... 17

Chapter 5: Continuous logic versus first-order logic .................................................. 53

Chapter 6: Conclusion .................................................................................................... 59

References .......................................................................................................................... 61
1. Introduction

Multi-valued logics are formal systems that deviate from classical logic by allowing for more than two truth values. An early version of multi-valued logic appeared in 1920 with the development of a three-valued logic by Łukasiewicz. This was extended to an infinite-valued propositional logic for which Pavelka proved a completeness result [3]. The next big step for multi-valued logics occurred in 1966, when Chang and Keisler generalized the set of truth values to be any compact Hausdorff space [5]. Among their results were suitable rephrasings of the compactness theorem and the Downward Löwenheim-Skolem theorem. Ultimately, their construction proved to be too general for applications to other areas of mathematics and little was done until the need arose in model theory to find a framework in which to study metric structures. Metric structures consist of complete, bounded metric spaces on which there are distinguished elements and uniformly continuous functions.

This paper presents a recent incarnation of multi-valued logic called continuous first-order logic, or simply continuous logic, introduced by Ben Yaacov and Usvyatsov in [4] to study metric structures. It is a special case of Chang and Keisler’s logic in [5] in the sense that the set of truth values is restricted to the unit interval [0, 1]. However, it differs from [5] by considering structures equipped with a distinguished metric instead of the traditional equality relation. In the same way that the equality symbol = in first-order logic is always interpreted to be real equality, the language of continuous logic contains a distinguished binary predicate $d$ that is always interpreted to be the metric. First-order model theory coincides with continuous model theory if we equip the underlying set of any first-order
model with the discrete metric. Thus continuous logic is an extension of first-order logic in the sense that it allows more structures to be models; see, e.g., Iovino [11]. These structures include Hilbert spaces, Banach spaces, and probability algebras. Probability algebras, which arise from taking quotients of probability spaces modulo their null sets and become complete metric spaces under \( d(a, b) = \mu(a \triangle b) \), where \( \mu \) is the probability measure, will be the focus of the first part of this paper.

Although continuous logic was developed by model theorists for model theoretic uses, some research about its properties as a logic has also been done. Notably, Ben Yaacov and Pederson formulated in [3] a set of axioms that yields a version of the completeness theorem. A comparison of the completeness theorem and other aspects of continuous logic to their counterparts in first-order logic will be presented in the second part of this paper.
2. Preliminaries

Let \((M, d)\) be a complete metric space with diameter 1.

**Definition.** A *predicate* on \(M\) is a uniformly continuous function \(P : M^n \to [0, 1]\). A *function* on \(M\) is a uniformly continuous function \(f : M^n \to M\).

**Definition.** A *metric structure* \(\mathcal{M}\) consists of a complete metric space \((M, d)\) with diameter 1, a set \(\{P_i : i \in I\}\) of predicates on \(M\), a set \(\{f_j : j \in J\}\) of functions on \(M\), and a set \(\{c_k : k \in K\}\) of distinguished elements of \(M\). We denote it by

\[
\mathcal{M} = (M, P_i, f_j, c_k : i \in I, j \in J, k \in K).
\]

**Convention.** We will regard constants as 0-ary functions and suppress explicit mention of them in future definitions. We treat \(d\) as a binary predicate since it is a uniformly continuous function from \(M^2 \to [0, 1]\); this is proven in 4.9. The set of natural numbers \(\{0, 1, \ldots\}\) is denoted by \(\mathbb{N}\).

Here are some examples of metric structures.

1. Let \((M, d)\) be any complete metric space of diameter 1. Then \(\mathcal{M} = (M, d)\) is a metric structure with no functions and whose sole predicate is the binary predicate \(d\).

2. Let \(\mathcal{M} = (M, P_i, f_j, c_k : i \in I, j \in J, k \in K)\) be a structure in first-order logic. Equip \(M\) with the discrete metric: for all \(x, y \in M\), \(d_M(x, y) = 1\) if and only if \(x \neq y\). This makes \(M\) a complete metric space with diameter 1. Predicates take
values in \( \{0, 1\} \subset [0, 1] \). Since the metric is discrete, all predicates and functions are trivially uniformly continuous. Under the intended interpretation of 0 as \( \top \) and 1 as \( \bot \), \( d \) becomes a generalization of equality: \( x = y \) is true in \( M \) if and only if \( d_M(x, y) = 0 \). In this sense, first-order models are a special case of metric structures.

(3) Given a probability space \((X, \mathcal{B}, \mu)\), we can construct a probability structure, a metric structure based on the metric space \((M, d)\) where \( M \) consists of elements of \( \mathcal{B} \) modulo 0 and \( d \) is the measure of symmetric difference. We take \( \mu \) to be a unary predicate and the Boolean operations \( \cup, \cap, \cdot \) to be functions on \( M \). We will study an axiomatization of this class of structures in the fourth chapter of this paper.

**Definition.** Let \((M, d), (M', d')\) be metric spaces. Let \( s : M \to M' \) be a function. \( \delta_s : (0, 1] \to (0, 1] \) is a modulus of uniform continuity for \( s \) if for all \( \epsilon \in (0, 1] \) and \( x, y \in M \), we have \( d(x, y) < \delta_s(\epsilon) \implies d'(s(x), s(y)) < \epsilon \).

Note that \( s \) is uniformly continuous if and only if it has a modulus of uniform continuity.

**Convention.** If \((M, d)\) is a metric space and \( f : M^n \to M \) is an \( n \)-ary function, then we always equip \( M^n \) with the maximum metric: given \( x = (x_1, ..., x_n), y = (y_1, ..., y_n) \in M^n \), we use \( d_{\text{max}}(x, y) = \max\{d(x_i, y_i) : i = 1, ..., n\} \).

If \((M, d)\) is a metric space and \( P : M^n \to [0, 1] \) is a \( n \)-ary predicate, then we always equip \([0, 1]\) with the usual metric: given \( a, b \in [0, 1] \), \( d_{[0,1]}(a, b) = |a - b| \).

**Definition.** A continuous signature is a quadruple \( L = (\mathcal{R}, \mathcal{F}, \eta, \mathcal{G}) \) such that

1. \( \mathcal{R} \) contains a distinguished binary predicate \( d \)
2. \( \mathcal{R} \cap \mathcal{F} = \emptyset \)
\( \eta : \mathcal{R} \cup \mathcal{F} \to \mathbb{N} \)

\( \mathcal{G} = \{ \delta_s : (0, 1] \to (0, 1] : s \in \mathcal{R} \cup \mathcal{F} \} \)

\( \mathcal{R} \) is the set of relation symbols, \( \mathcal{F} \) is the set of function symbols, and \((\mathcal{R} \setminus \{d\}) \cup \mathcal{F}\) is the set of nonlogical symbols of \( L \). The arity of a relation or function symbol \( s \) is given by \( \eta(s) \).

\( \mathcal{G} \) is the set of moduli of uniform continuity for each predicate and function symbol. The moduli of uniform continuity are not syntactic objects; rather, they are functions \((0, 1] \to (0, 1]\) fixed by the signature. The cardinality of \( L \), denoted by \(|L|\), is the cardinality of the set of nonlogical symbols of \( L \).

The logical symbols of any signature \( L \) consist of the following: a binary connective \( \land \), two unary connectives, \( \neg \) and \( \frac{1}{2} \), two quantifiers, sup and inf, countably many variables \( V = \{v_0, v_1, \ldots\} \). Since we always interpret \( d \) to be the metric, \( d \) is also considered a logical symbol. This parallels the treatment of the equality symbol in first-order logic.

2.1. **Definition.** Let \( L = (\mathcal{R}, \mathcal{F}, \eta, \mathcal{G}) \) be a continuous signature. An \( L \)-structure is an ordered pair \( \mathcal{M} = (M, \rho) \) where \( M \) is a set and \( \rho \) is a function on \( \mathcal{R} \cup \mathcal{F} \) such that

1. for each \( f \in \mathcal{F} \), \( \rho \) assigns a map \( f^\mathcal{M} : M^{\eta(f)} \to M \)
2. for each \( P \in \mathcal{R} \), \( \rho \) assigns a map \( P^\mathcal{M} : M^{\eta(P)} \to [0, 1] \)
   (in particular, \( \rho \) assigns \( d \) to a complete metric \( d^\mathcal{M} : M \times M \to [0, 1]\))
3. for each \( f \in \mathcal{F} \), \( \delta_f \) is a modulus of uniform continuity for \( f \): for all \( \epsilon \in (0, 1] \) and \( x = (x_1, \ldots, x_{\eta(f)}) \), \( y = (y_1, \ldots, y_{\eta(f)}) \in M^{\eta(f)} \),
   \[
   d_{\max}(x, y) < \delta_f(\epsilon) \implies d^\mathcal{M}(f(x), f(y)) < \epsilon
   \]
(4) for each $P \in \mathcal{R}$, $\delta_P$ is a modulus of uniform continuity for $P$: for all $\epsilon \in (0, 1]$ and $x = (x_1, ..., x_{\eta(P)})$, $y = (y_1, ..., y_{\eta(P)}) \in M^{\eta(P)}$, $d_{\max}(x, y) < \delta_P(\epsilon) \implies |P(x) - P(y)| < \epsilon$

**Remark.** Part (2) of the above definition establishes the set of truth values as $[0,1]$.

**Definition.** Fix a continuous signature $L = (\mathcal{R}, \mathcal{F}, \eta, \mathcal{G})$. $L$-terms are defined inductively. Variables and constant symbols are $L$-terms. If $f \in \mathcal{F}$ is an $n$-ary function symbol and $t_1, ..., t_n$ are $L$-terms, then $f(t_1, ..., t_n)$ is an $L$-term. Atomic $L$-formulas are expressions of the form $P(t_1, ..., t_n)$, where $P \in \mathcal{R}$ is an $n$-ary predicate symbol and $t_1, ..., t_n$ are $L$-terms. The class of $L$-formulas is the smallest class of expressions that contains all atomic formulas, is closed under $\land$, $\lor$, and $\frac{1}{2}$, and satisfies the following condition: If $\phi$ is an $L$-formula and $x$ is a variable, then $\sup_x \phi$ is an $L$-formula.

2.2. **Definition.** Let $\mathcal{M} = (M, \rho)$ be an $L$-structure. An $\mathcal{M}$-assignment is a function $\sigma : V \to M$. Given an $\mathcal{M}$-assignment $\sigma$, $x \in V$, and $a \in M$, define the $\mathcal{M}$-assignment $\sigma^a_x$ as follows: for all $y \in V$,

$$
\sigma^a_x(y) := \begin{cases} 
a & \text{if } x = y \\
\sigma(y) & \text{otherwise}
\end{cases}
$$

Fix an $\mathcal{M}$-assignment $\sigma$. Define $t^{\mathcal{M}, \sigma}$, the interpretation under $\sigma$ of a term $t$ in $\mathcal{M}$, inductively in the same way as in first-order logic. Let $\phi$ be a formula. Define $\mathcal{M}(\phi, \sigma)$, the value of $\phi$ in $\mathcal{M}$ under $\sigma$, by induction on formulas:

1. $\mathcal{M}(P(t_1, ..., t_{\eta(P)}), \sigma) := P^{\mathcal{M}}(t_1^{\mathcal{M}, \sigma}, ..., t_{\eta(P)}^{\mathcal{M}, \sigma})$

2. $\mathcal{M}(\alpha \lor \beta, \sigma) := \max(\mathcal{M}(\alpha, \sigma) - \mathcal{M}(\beta, \sigma), 0)$

3. $\mathcal{M}(\neg \alpha, \sigma) := 1 - \mathcal{M}(\alpha, \sigma)$
(4) $\mathcal{M}(\frac{1}{2}\alpha, \sigma) := \frac{1}{2} \mathcal{M}(\alpha, \sigma)$

(5) $\mathcal{M}(\sup_{x}\alpha, \sigma) := \sup\{\mathcal{M}(\alpha, \sigma^a) : a \in M\}$

When $\phi$ is a sentence, we write $\mathcal{M}(\phi)$ instead of $\mathcal{M}(\phi, \sigma)$.

2.3. We introduce the following abbreviations for use in future examples. Let 1 stand for $\neg(\phi \div \phi)$ for any sentence $\phi$. Let $(\frac{1}{2})^n$ stand for

$$\underbrace{\frac{1}{2} \cdot \frac{1}{2} \cdots} \frac{1}{2} 1, \text{ n times}.$$  

This notation is natural because for any formula $\phi$, the truth value of $\neg(\phi \div \phi)$ in any model under any assignment is 1:

$$\mathcal{M}(\neg(\phi \div \phi), \sigma) = 1 - \mathcal{M}(\phi \div \phi, \sigma) = 1 - \max(\mathcal{M}(\phi, \sigma) - \mathcal{M}(\phi, \sigma), 0) = 1.$$ 

Similarly, for any $n$, the truth value of $(\frac{1}{2})^n$ is always $(\frac{1}{2})^n$. We can use this to construct a sentence $\alpha$ whose truth value in any model $\mathcal{M}$ is $\frac{3}{4}$. First, define a binary connective $+_{0}$ by $\phi +_{0} \psi = \neg(\psi \rightarrow \neg\phi)$. Then

$$\mathcal{M}(\phi +_{0} \psi, \sigma) = \min\{\mathcal{M}(\phi, \sigma) + \mathcal{M}(\psi, \sigma), 1\}$$

for all models $\mathcal{M}$ and assignments $\sigma$. Note that $+_{0}$ is classically equivalent to conjunction.

Now take $\alpha$ to be $\frac{1}{2} +_{0} (\frac{1}{2})^2$. Then $\mathcal{M}(\alpha, \sigma) = \min\{\frac{1}{2} + \frac{1}{4}, 1\} = \frac{3}{4}$. In this way, for every dyadic number $\frac{k}{2^n}$ where $k, n \in \mathbb{N}, k \leq 2^n$, we can produce a sentence whose truth value is $\frac{k}{2^n}$, so the set of truth values generated by $\{-, \div, \frac{1}{2}\}$ from 1 is $\mathbb{D} = \{\frac{k}{2^n} : k, n \in \mathbb{N}, k \leq 2^n\}$.

In propositional logic, a chosen set of connectives must be functionally complete; every Boolean function should be representable in terms of a propositional formula formed from these connectives. The dual notion of functional completeness in continuous logic is $fullness$. In general, for each $n \geq 1$, an $n$-ary connective $f$ is a continuous function from
A system of connectives is a family $\mathcal{F} = (F_n \mid n \geq 1)$ where each $F_n$ is a set of $n$-ary connectives. $\mathcal{F}$ is closed if the following hold:

1. For all $n$, $F_n$ contains the projection $(a_1, \ldots, a_n) \mapsto a_j$ for each $j = 1, \ldots, n$.
2. For all $m, n$, if $u \in F_n$ and $v_1, \ldots, v_n \in F_m$, then $w : [0, 1]^m \to [0, 1]$ defined by $w(\overline{a}) = u(v_1(\overline{a}), \ldots, v_n(\overline{a}))$ is in $F_m$.

For each system of connectives $\mathcal{F}$, let $\overline{\mathcal{F}}$ denote the smallest closed system of connectives containing $\mathcal{F}$. $\mathcal{F}$ is full if for each $n$, $\overline{\mathcal{F}}$ is dense in the set $C_n$ of all $n$-ary connectives $[0, 1]^n \to [0, 1]$ with respect to the topology on $C_n$ obtained from the metric $d(f, g) = \sup_{\overline{x} \in [0, 1]^n} |f(\overline{x}) - g(\overline{x})|$. In other words, for all $n \geq 1$, $\epsilon > 0$, and $f \in C_n$, there exists $g \in \overline{\mathcal{F}}$ such that for all $\overline{x} \in [0, 1]^n$, $|f(\overline{x}) - g(\overline{x})| < \epsilon$.

Instead of adding a symbol for each connective, which would render the language uncountable, we may choose any full system of connectives, allowing us to approximate any connective to arbitrary precision.

**Fact.** (6.6 in [2]) The system of connectives $\{\lor, \neg, \frac{1}{2}\}$ is full.

Thus we choose $\{\lor, \neg, \frac{1}{2}\}$, which is the analogue of $\{\to, \neg\}$, a popular choice of functionally complete connectives in first-order logic. Interestingly, although $\{\neg, \land, \lor\}$ is functionally complete in first-order logic, it is not full in continuous logic with $\land$ and $\lor$ as defined in 3.3; by 1.7 in [4], all $n$-ary connectives constructed using $\neg, \land, \lor$ are 1-Lipschitz in each argument, but $\frac{1}{2}$ is 2-Lipschitz, so $\frac{1}{2}$ cannot be constructed from $\neg, \land, \lor$.

### 2.4. Definition

Let $\mathcal{M}$ be an $L$-structure, $\sigma$ an assignment, and $\Sigma$ a set of $L$-formulas. $(\mathcal{M}, \sigma)$ is a model of $\Sigma$, written $(\mathcal{M}, \sigma) \models \Sigma$, if for all $\phi \in \Sigma$, $\mathcal{M}(\phi, \sigma) = 0$. We will abbreviate $(\mathcal{M}, \sigma) \models \{\phi\}$ by $(\mathcal{M}, \sigma) \models \phi$. $\Sigma$ is satisfiable if it has a model. We write $\Sigma \models \phi$ if every model of $\Sigma$ is a model of $\phi$. $\phi$ is valid if $\emptyset \models \phi$. 
3. Axioms of Continuous Logic

Definition. Given a term $t$, a variable $x$, and a formula $\alpha$, the substitution of $t$ for $x$ in $\alpha$, denoted $\alpha[t/x]$, is the result of replacing every free occurrence of $x$ in $\alpha$ by $t$.

Define $\phi \rightarrow \psi := \psi \land \phi$. The following axiom system from [3] for continuous logic allows us to prove the completeness result in 5.3. The axioms of continuous logic are the universal closures of:

1. $\alpha \rightarrow (\beta \rightarrow \alpha)$
2. $(\alpha \rightarrow \beta) \rightarrow ((\beta \rightarrow \gamma) \rightarrow (\alpha \rightarrow \gamma))$
3. $((\alpha \rightarrow \beta) \rightarrow \beta) \rightarrow ((\beta \rightarrow \alpha) \rightarrow \alpha)$
4. $(\neg \alpha \rightarrow \neg \beta) \rightarrow (\beta \rightarrow \alpha)$
5. $(\frac{1}{2} \alpha \rightarrow \alpha) \rightarrow \frac{1}{2} \alpha$
6. $\frac{1}{2} \alpha \rightarrow (\frac{1}{2} \alpha \rightarrow \alpha)$
7. $\sup_x (\alpha \rightarrow \beta) \rightarrow (\sup_x \alpha \rightarrow \sup_x \beta)$
8. $\sup_x \alpha \rightarrow \alpha[t/x]$, if no variable in $t$ is bound by a quantifier in $\alpha$
9. $\alpha \rightarrow \sup_x \alpha$, if $x$ is not free in $\alpha$
10. $d(x, x)$
11. $d(x, y) \rightarrow d(y, x)$
12. $d(y, z) \rightarrow (d(x, y) \rightarrow d(x, z))$
13. For each $f \in F$, $\epsilon \in (0, 1]$, and $r, q \in \mathbb{D}$, if $r > \epsilon$ and $q < \delta_f(\epsilon)$, then

$$\left( \bigvee_{i=1, \ldots, \eta(f)} d(x_i, y_i) \rightarrow q \right) \lor (r \rightarrow d(f(x_1, \ldots, x_{\eta(f)}), f(y_1, \ldots, y_{\eta(f)})))$$
(14) For each $P \in \mathcal{R}$, $\epsilon \in (0, 1]$, and $r, q \in \mathbb{D}$, if $r > \epsilon$ and $q < \delta_P(\epsilon)$, then

$$
\left( \bigvee_{i=1, \ldots, \eta(P)} d(x_i, y_i) \rightarrow q \right) \lor (r \rightarrow |P(x_1, \ldots, x_{\eta(P)}) - P(y_1, \ldots, y_{\eta(P)})|)
$$

The only rule of inference in continuous logic is modus ponens

$$\begin{array}{c}
\alpha, \alpha \rightarrow \beta \\
\hline
\beta
\end{array}$$

Formal deductions and the provability relation $\vdash$ are defined in the usual way.

3.1. Interpreting implication

Let $\mathcal{M}$ be an $L$-structure, $\sigma$ an assignment, and $\phi, \psi$ $L$-formulas. By definition, $\mathcal{M}(\phi \rightarrow \psi, \sigma) = \mathcal{M}(\psi \rightarrow \phi, \sigma) = \max(\mathcal{M}(\psi, \sigma) - \mathcal{M}(\phi, \sigma), 0)$, so

$$(\mathcal{M}, \sigma) \models \phi \rightarrow \psi \iff \mathcal{M}(\psi, \sigma) \leq \mathcal{M}(\phi, \sigma).$$

In words, $\phi \rightarrow \psi$ holds in $(\mathcal{M}, \sigma)$ if and only if $\psi$ is at most as true as $\phi$. If we restrict the set of truth values to $\{0, 1\}$, then this says that the only instance where $\phi \rightarrow \psi$ does not hold in $(\mathcal{M}, \sigma)$ is when $\mathcal{M}(\psi, \sigma) > \mathcal{M}(\phi, \sigma)$, i.e. when $\mathcal{M}(\psi, \sigma) = 1$ and $\mathcal{M}(\phi, \sigma) = 0$.

An interesting phenomenon occurs when $\psi$ is $\neg \phi$. In this case,

$$(\mathcal{M}, \sigma) \models \phi \rightarrow \neg \phi \iff 1 - \mathcal{M}(\phi, \sigma) = \mathcal{M}(\neg \phi, \sigma) \leq \mathcal{M}(\phi, \sigma) \iff \mathcal{M}(\phi, \sigma) \geq \frac{1}{2}.$$ 

Thus for a formula $\phi$ to imply its own negation in $(\mathcal{M}, \sigma)$, it only needs to be more false than $\frac{1}{2}$ in $(\mathcal{M}, \sigma)$.

3.2. Defining disjunction
In first-order logic, there are many ways to define disjunction in terms of \( \to \) and \( \neg \). For example, we could take \( \phi \lor_1 \psi := \neg \psi \to \phi \) and \( \phi \lor_2 \psi := (\psi \to \phi) \to \phi \). With the semantics of continuous logic,

\[
\mathcal{M}(\phi \lor_1 \psi, \sigma) = \mathcal{M}(\neg \psi \to \phi, \sigma) = \max(\mathcal{M}(\psi, \sigma) + \mathcal{M}(\phi, \sigma) - 1, 0)
\]

\[
\mathcal{M}(\phi \lor_2 \psi, \sigma) = \mathcal{M}((\psi \to \phi) \to \phi, \sigma) = \min(\mathcal{M}(\phi, \sigma), \mathcal{M}(\phi, \sigma))
\]

If we restrict the range of \( \mathcal{M}(\cdot, \sigma) \) to \{0, 1\}, then \( \mathcal{M}(\cdot \lor_1 \cdot, \sigma) \equiv \mathcal{M}(\cdot \lor_2 \cdot, \sigma) \), reflecting the fact that \( \lor_1 \) and \( \lor_2 \) are equivalent in first-order logic. To see that they are not equivalent in continuous logic, consider \( \phi = \frac{1}{2} \) as defined in 2.3. Since \( \mathcal{M}(\phi, \sigma) = \frac{1}{2} \),

\[
\mathcal{M}(\phi \lor_1 \phi, \sigma) = \max(\mathcal{M}(\phi, \sigma) + \mathcal{M}(\phi, \sigma) - 1, 0) = \max(\frac{1}{2} + \frac{1}{2} - 1, 0) = 0
\]

\[
\mathcal{M}(\phi \lor_2 \phi, \sigma) = \min(\mathcal{M}(\phi, \sigma), \mathcal{M}(\phi, \sigma)) = \min(\frac{1}{2}, \frac{1}{2}) = \frac{1}{2}.
\]

Our choice of disjunction \( \lor \) should give the same truth value to \( \phi \) and to \( \phi \lor \phi \). Furthermore, for all models \( \mathcal{M} \) and assignments \( \sigma \), it should satisfy

\[
(\mathcal{M}, \sigma) \models \phi \lor \psi \iff (\mathcal{M}, \sigma) \models \phi \text{ or } \mathcal{M} \models \psi
\]

and \( \mathcal{M}(\phi \lor \psi, \sigma) = \min(\mathcal{M}(\phi, \sigma), \mathcal{M}(\psi, \sigma)) \), so the most natural choice is \( \lor = \lor_2 \) (or anything equivalent to it in continuous logic).

Similarly, conjunction \( \land \) should satisfy

\[
(\mathcal{M}, \sigma) \models \phi \land \psi \iff (\mathcal{M}, \sigma) \models \phi \text{ and } \mathcal{M} \models \psi
\]

and \( \mathcal{M}(\phi \land \psi, \sigma) = \max(\mathcal{M}(\phi, \sigma), \mathcal{M}(\psi, \sigma)) \), both of which are achieved by taking \( \phi \land \psi := \neg(\neg \phi \lor \neg \psi) \). These remarks are summarized below.
3.3. We introduce the following abbreviations:

**disjunction** \( \phi \lor \psi := (\psi \rightarrow \phi) \rightarrow \phi \)**

**conjunction** \( \phi \land \psi := \neg(\neg \phi \lor \neg \psi) \)**

**logical distance** \( |\phi - \psi| := (\phi \rightarrow \psi) \land (\psi \rightarrow \phi) \)

\( \inf_x \phi := \neg \sup_x \neg \phi \)

Under the semantics of continuous logic,

1. \( M(\phi \rightarrow \psi, \sigma) = \max(M(\psi, \sigma) - M(\phi, \sigma), 0) \)
2. \( M(\phi \lor \psi, \sigma) = \min(M(\phi, \sigma), M(\psi, \sigma)) \)
3. \( M(\phi \land \psi, \sigma) = \max(M(\phi, \sigma), M(\psi, \sigma)) \)
4. \( M(|\phi - \psi|, \sigma) = |M(\phi, \sigma) - M(\psi, \sigma)| \) (one-dimensional Euclidean distance)
5. \( M(\inf_x \phi, \sigma) = \inf\{M(\phi, \sigma_x^a) : a \in M\} \)

If we restrict the truth values to \( \{0, 1\} \), the interpretations of sup, inf, \( \lor, \land, \) and \( \rightarrow \) coincide with those of \( \forall, \exists, \) disjunction, conjunction, and implication, respectively, in first-order logic. For example, \( (M, \sigma) \models \inf_x \phi \iff M(\inf_x \phi, \sigma) = 0 \iff \) there exists \( a \in M \) such that \( M(\phi, \sigma_x^a) = 0 \). When it is easier to do so, we will write \( \phi \leftrightarrow \psi \) instead of \( |\phi - \psi| \).

**Remark.** Under the abbreviations in 3.3, axiom 3 states that \( \lor \) is commutative. Axioms 1–4 are the axioms for Łukasiewicz propositional logic [3]. Their relation to the axioms of classical propositional logic is explored in 5.5.
Axioms 5 and 6 establish the behavior of $\frac{1}{2}$. Recall from 2.3 that $\alpha +_0 \beta := \neg(\beta \rightarrow \neg\alpha)$ and $M(\alpha +_0 \beta, \sigma) = \min(M(\alpha, \sigma) + M(\beta, \sigma), 1)$ for any $L$-structure $M$ and $M$-assignment $\sigma$. In particular,

$$M(\frac{1}{2}\alpha + \frac{1}{2}\alpha, \sigma) = \min(M(\frac{1}{2}\alpha, \sigma) + M(\frac{1}{2}\alpha, \sigma), 1)$$

$$= \min(\frac{1}{2}M(\alpha, \sigma) + \frac{1}{2}M(\alpha, \sigma), 1)$$

$$= \min(M(\alpha, \sigma), 1).$$

Axioms 7–9 establish the behavior of sup and are direct translations of the quantifier axioms in first-order logic [7].

Axioms 10–14 are translations of properties of $=$ in first-order logic. Recall the equality axioms of first-order logic (from [1]):

(i) reflexivity: $x = x$

(ii) substitutivity: $x = y \rightarrow (\phi \rightarrow \phi')$ for any formula $\phi$, where $\phi'$ is obtained from $\phi$ by replacing some free occurrences of $x$ with $y$.

Let $M$ be an $L$-structure and let $\sigma$ be an assignment. If we replace $a = b$ by $d(a, b)$, then axiom 10 is a direct translation of (i). Consequences of (ii) include symmetry, $(x = y) \rightarrow (y = x)$, and transitivity, $(x = y) \rightarrow ((x = z) \rightarrow (y = z))$, which are encoded in axioms 11 and 12, respectively. So the property of $=$ being an equivalence relation translates, via axioms 10–12, into $d$ being a pseudo-metric. The reason why we do not need to specify that $d$ is a genuine complete metric is given in 3.4.

Equality axiom (ii) also makes $=$ a congruence relation; for each predicate $P$, $(x = y) \rightarrow (P(x, z) \rightarrow P(y, z))$. By 3.1,
\[(\mathcal{M}, \sigma) \models (d(x, y)) \rightarrow (P(x, z) \rightarrow P(y, z))\]

\[
\iff \mathcal{M}(P(x, z) \rightarrow P(y, z), \sigma) \leq \mathcal{M}(d(x, y), \sigma)
\]

\[
\iff \max(\mathcal{M}(P(y, z), \sigma) - \mathcal{M}(P(x, z), \sigma), 0) \leq d_M(x, y).
\]

We also have \(\max(\mathcal{M}(P(x, z), \sigma) - \mathcal{M}(P(y, z), \sigma), 0) \leq d_M(y, x) = d_M(x, y),\) so

\[
|P^\mathcal{M}(x, z) - P^\mathcal{M}(y, z)| = |\mathcal{M}(P(x, z), \sigma) - \mathcal{M}(P(y, z), \sigma)| \leq d_M(x, y).
\]

In other words, predicates should be 1-Lipschitz with respect to \(d,\) although 1-Lipschitz is relaxed to uniformly continuous. A similar argument using \((x = y) \rightarrow (f(x, z) = f(y, z))\) translated to \(d(x, y) \rightarrow (d(f(x, z), c) \rightarrow d(f(y, z), c))\) shows that functions should be 1-Lipschitz, which is likewise relaxed to uniformly continuous. To express these requirements, axioms 13 and 14 define the moduli of uniform continuity for each function and predicate symbol.

3.4. We now introduce a more general notion of metric structure and summarize its relation to metric structures.

**Definition.** Using the notation from 2.1, an \(L\)-pre-structure is an \(L\)-structure such that \(\rho\) assigns \(d\) to a pseudo-metric.

**Definition.** Let \(L = (\mathcal{R}, \mathcal{F}, \eta, \mathcal{G})\) be a continuous signature. Let \(\mathcal{M}\) and \(\mathcal{N}\) be \(L\)-pre-structures with underlying sets \(M\) and \(N\), respectively. A function \(h : M \rightarrow N\) is an elementary \(L\)-morphism if it satisfies the following conditions:

1. for all \(f \in \mathcal{F}\) and \(a_1, \ldots, a_{\eta(f)} \in M,\)

\[
h(f^\mathcal{M}(a_1, \ldots, a_{\eta(f)})) = f^\mathcal{N}(h(a_1), \ldots, h(a_{\eta(f)})).
\]
(2) for all $P \in \mathcal{R}$ and $a_1, \ldots, a_{\eta(P)} \in M$,

$$P^M(a_1, \ldots, a_{\eta(P)}) = P^N(h(a_1), \ldots, h(a_{\eta(P)})).$$

(3) for all $\mathcal{M}$-assignments $\sigma$ and formulas $\phi$, $\mathcal{M}(\phi, \sigma) = \mathcal{N}(\phi, h \circ \sigma)$.

The definitions of assignment, interpretation, and value in $L$-pre-structures are the same as those in $L$-structures from 2.2. Given an $L$-pre-structure $\mathcal{M}$, an assignment $\sigma$, and a set of $L$-formulas $\Sigma$, we say that $(\mathcal{M}, \sigma)$ is a pre-model of $\Sigma$ if for all $\phi \in \Sigma$, $\mathcal{M}(\phi, \sigma) = 0$. We say that a set of formulas $\Sigma$ is pre-satisfiable if it has a pre-model.

**Theorem.** (6.9 from [3]) Let $L$ be a continuous signature and let $\mathcal{M}$ be an $L$-pre-structure. Then there exists an $L$-structure $\widehat{\mathcal{M}}$ and an elementary $L$-morphism of $\mathcal{M}$ into $\widehat{\mathcal{M}}$.

$\widehat{\mathcal{M}}$ can be constructed by taking the quotient space of $\mathcal{M}$ under the equivalence relation $x \sim y \iff d(x, y) = 0$.

**Corollary.** A set of formulas $\Gamma$ is satisfiable if and only if $\Gamma$ is pre-satisfiable.

Thus $L$-pre-structures and $L$-structures are indistinguishable in continuous logic. The analogue of this fact in first-order logic is that $=$ is a congruence relation, so the quotient of a first-order structure $\mathcal{K}$ by $=^\mathcal{K}$ is indistinguishable from $\mathcal{K}$ in first-order logic.

We end this section with definitions of elementary equivalence and elementary substructure. They coincide with the definitions from first-order logic when the set of truth values is restricted to $\{0, 1\}$.

**Definition.** Let $\mathcal{M}$ and $\mathcal{N}$ be $L$-structures. $\mathcal{M}$ and $\mathcal{N}$ are elementarily equivalent if $\mathcal{M}(\phi) = \mathcal{N}(\phi)$ for all $L$-sentences $\phi$. $\mathcal{M}$ is an elementary substructure of $\mathcal{N}$ if $M \subseteq N$. 
and for all formulas $\phi(x_1, \ldots, x_n)$ and $a_1, \ldots, a_n \in M$, $\mathcal{M}(\phi, \sigma) = \mathcal{N}(\phi, \sigma')$ where $\sigma$ is any $\mathcal{M}$-assignment and $\sigma'$ is any $\mathcal{N}$-assignment such that $\sigma(x_i) = \sigma'(x_i) = a_i$ for each $i = 1, \ldots, n$.

**Definition.** Let $L$ be a continuous signature. A *sentence* is an $L$-formula with no free variables. A *theory* $T$ is a set of sentences. The theory of an $L$-structure $\mathcal{M}$, denoted by $\text{Th}(\mathcal{M})$, is the set of sentences true in $\mathcal{M}$.

**Remark.** If $\mathcal{M}$ and $\mathcal{N}$ are elementarily equivalent, then $\text{Th}(\mathcal{M}) = \text{Th}(\mathcal{N})$. The converse also holds: Suppose for $L$-structures $\mathcal{M}$ and $\mathcal{N}$, $\text{Th}(\mathcal{M}) = \text{Th}(\mathcal{N})$. Let $\phi$ be a sentence and let $q = \mathcal{M}(\phi)$. By 2.3, there exists an $L$-sentence $\psi$ such that in all $L$-structures $\mathcal{Q}$, $\mathcal{Q}(\psi) = q$. In particular $\mathcal{M}(\psi) = q$, so $\mathcal{M}(\phi \leftrightarrow \psi) = 0$. Since $\text{Th}(\mathcal{M}) = \text{Th}(\mathcal{N})$, $\mathcal{N}(\phi \leftrightarrow \psi) = 0$. Since $\mathcal{N}(\psi) = q$, we have $\mathcal{N}(\phi) = \mathcal{N}(\psi) = q = \mathcal{M}(\phi)$, so $\mathcal{M}$ and $\mathcal{N}$ are elementarily equivalent.
4. Probability Algebras

**Definition.** [10] A *Boolean algebra* is a set $A$ that includes two distinguished constants 0 and 1 and on which there are defined two binary operations $\land$ and $\lor$, a unary operation $\cdot^c$ which satisfy the following rules:

1. $\forall x \; x \lor x = x$
2. $\forall x \; x \land x = x$
3. $\forall x \forall y \; x \lor y = y \lor x$
4. $\forall x \forall y \; x \land y = y \land x$
5. $\forall x \forall y \; (x \land y) \lor y = y$
6. $\forall x \forall y \; (x \lor y) \land y = y$
7. $\forall x \forall y \forall z \; (x \lor y) \lor z = x \lor (y \lor z)$
8. $\forall x \forall y \forall z \; (x \land y) \land z = x \land (y \land z)$
9. $\forall x \forall y \forall z \; x \lor (y \land z) = (x \lor y) \land (x \lor z)$
10. $\forall x \forall y \forall z \; x \land (y \lor z) = (x \land y) \lor (x \land z)$
11. $\forall x \; x \lor x^c = 1$
12. $\forall x \; x \land x^c = 0$
13. $0 \neq 1$

A Boolean algebra becomes partially ordered if we define $x \leq y$ to be $x \land y = x$ (reflexive by (2), antisymmetric by (4), transitive by (8)). A Boolean algebra $A$ is a
**Boolean σ-algebra** if for each sequence \( (x_n) \) of elements of \( \mathcal{A} \) there is a least element \( x \in \mathcal{A} \) such that \( x_n \leq x \) for all \( n \). We denote this least upper bound by \( \sup_n x_n \).

4.1. **Facts about Boolean algebras.** Let \( \mathcal{B} \) be a Boolean algebra. For all \( a, b \in \mathcal{B} \), let 

\[
a \setminus b := a \wedge b^c.
\]

4.1.1. For all \( a \in \mathcal{B} \), \( a \lor 0 = a \), \( a \wedge 0 = 0 \), \( a \lor 1 = 1 \), and \( a \wedge 1 = a \).

4.1.2. **Disjoint unions.** For all \( a_0 \leq a_1 \leq a_2 \in \mathcal{B} \), \( a_2 \setminus a_0 = (a_2 \setminus a_1) \lor (a_1 \setminus a_0) \) and 

\[
(a_2 \setminus a_1) \land (a_1 \setminus a_0) = 0.
\]

Furthermore, \( i < j \in \{0, 1, 2\} \) implies \( a_i \setminus a_j = 0 \).

**Proof.** Observe 

\[
(a_2 \land a_1^c) \lor (a_1 \land a_0^c) = (a_2 \land a_1^c) \lor ((a_2 \land a_1) \land a_0^c) = (a_2 \land a_1^c) \lor (a_2 \land (a_1 \land a_0^c)) = a_2 \land (a_1^c \lor a_1) \land (a_1^c \lor a_0^c) = a_2 \land (a_1^c \lor a_0^c) = a_2 \land a_0^c = a_2 \setminus a_0
\]

and 

\[
(a_2 \setminus a_1) \land (a_1 \setminus a_0) = (a_2 \setminus a_1^c) \land (a_1 \setminus a_0^c) = a_2 \setminus a_1 \land a_1 \land a_0 = 0.
\]

Now let \( i < j \in \{0, 1, 2\} \). Then \( a_i \setminus a_j = a_i \land a_j^c = (a_i \land a_j) \land a_j^c = a_i \land (a_j \land a_j^c) = a_i \land 0 = 0 \).

4.1.3. \( b \leq b_n \) implies \( b \lor b_n = b_n \).

**Proof.** Suppose \( b \leq b_n \). By definition \( b \land b_n = b \), so by axiom 5 \( b \lor b_n = (b \land b_n) \lor b_n = b_n \).

4.1.4. For all \( a, b, c \in \mathcal{B} \), \( (a \triangle b) \land c = (a \land c) \triangle (b \land c) \).

**Proof.** 

\[
(a \triangle b) \land c = ((a \land b^c) \lor (b \land a^c)) \land c = ((a \land b^c) \land c) \lor ((b \land a^c) \land c) = ((a \land c) \land b^c) \lor ((b \land c) \land a^c) = (((a \land c) \land b^c) \lor 0) \lor (((b \land c) \land a^c) \lor 0) = (((a \land c) \land b^c) \lor (a \land 0)) \lor (((b \land c) \land a^c) \lor (b \land c^c)) = \\
[((a \land c) \land b^c) \lor ((a \land c) \land c^c)] \lor [[[((b \land c) \land a^c) \lor (b \land (c \land c^c))] = \\
(((a \land c) \lor (a \land c^c)) \lor ((a \land c) \lor (b \land c^c)) = ((a \land c) \lor (b^c \lor c^c)) \lor \\
((a^c \lor c^c) \land (b \land c)) = ((a \land c) \land (b \land c)^c) \lor ((a \land c)^c \land (b \land c)) = (a \land c) \triangle (b \land c). \quad \Box
\]
4.1.5. For all \(a, b, c \in \mathcal{B}\), \(a \triangle c \leq (a \triangle b) \lor (b \triangle c)\).

**Proof.** Let \(a, b, c \in \mathcal{B}\). Then
\[
(a \land c^c) \lor ((a \lor b) \lor (b \lor c)) = (a \land c^c) \lor ((a \land b^c) \lor (a \land b^c)) \lor ((a \land b^c) \lor (b \land c^c)) \lor ((a \land b^c) \lor (b \land c^c)) \lor ((a \land b^c) \lor (b \land c^c)) = (a \land c^c) \lor (a \land b^c) \lor (a \land b^c) \lor (a \land b^c) \lor (a \land b^c) = (a \land c^c),
\]
so \(a \land c^c \leq (a \triangle b) \lor (b \triangle c)\). By symmetry of \(\triangle\), \(a^c \land c \leq (a \triangle b) \lor (b \triangle c)\), so \(a \triangle c \land ((a \triangle b) \lor (b \triangle c)) = ((a \land c^c) \lor (a^c \land c)) \land ((a \triangle b) \lor (b \triangle c)) = ((a \land c^c) \lor (a^c \land c) \lor (a \triangle b) \lor (b \triangle c)) = ((a \land c^c) \lor (a^c \land c) \lor (a \triangle b) \lor (b \triangle c)) = (a \land c^c) \lor (a^c \land c) = a \triangle c\). Hence, \(a \triangle c \leq (a \triangle b) \lor (b \triangle c)\).

\[
\square
\]

4.1.6. For all \(a, b \in \mathcal{B}\), if \(a \setminus b = 0\), then \(a \leq b\).

**Proof.** Let \(a, b \in \mathcal{B}\) and suppose \(a \setminus b = 0\). Then
\[
a = a \land 1 = a \land (b \lor b^c) = (a \land b) \lor (a \land b^c) = (a \land b) \lor (a \setminus b) = a \land b,
\]
so \(a \leq b\).

\[
\square
\]

4.1.7. For all \(a, b, c \in \mathcal{B}\), \((a \triangle b) \triangle c = a \triangle (b \triangle c)\)

**Definition.** A *probability space* is a triple \((X, \mathcal{B}, \mu)\) where \(X\) is a set, \(\mathcal{B}\) is a \(\sigma\)-algebra of subsets of \(X\), and \(\mu\) is a \(\sigma\)-additive measure on \(\mathcal{B}\) such that \(\mu(X) = 1\).

4.1. **Proposition.** Let \((X, \mathcal{B}, \mu)\) be a probability space. Define the relation \(\sim_\mu\) on \(\mathcal{B}\) by
\[
A \sim_\mu B \iff \mu(A \triangle B) = 0.
\]
Then \(\sim_\mu\) is an equivalence relation.

**Proof.** Clearly \(\sim_\mu\) is reflexive and symmetric. Let \(A, B, C \in \mathcal{B}\) such that \(A \sim_\mu B\) and \(B \sim_\mu C\). Then \(\mu(A \triangle B) = \mu(B \triangle C) = 0\). Since \(A \setminus B\) and \(B \setminus A\) are disjoint, \(0 = \mu(A \triangle B) = \mu((A \setminus B) \cup (B \setminus A)) = \mu(A \setminus B) + \mu(B \setminus A),\) so \(\mu(A \setminus B) = \mu(B \setminus A) = 0\).

Similarly, \(\mu(B \setminus C) = \mu(C \setminus B) = 0\). Since \(A \setminus C \subseteq (A \setminus B) \cup (B \setminus C)\) and \(C \setminus A \subseteq (C \setminus B) \cup (B \setminus A),\) \(\mu(A \triangle C) = \mu((A \setminus C) \cup (C \setminus A)) = \mu(A \setminus C) + \mu(C \setminus A) \leq \mu(A \setminus B) + \mu(B \setminus A) = 0\).
\[ \mu(A \setminus B) + \mu(B \setminus C) + \mu(C \setminus B) + \mu(B \setminus A) = 0. \] Thus \( A \sim_{\mu} C \), so \( \sim_{\mu} \) is transitive, and hence an equivalence relation.

**Notation.** Denote the equivalence class of \( A \in \mathcal{B} \) by \([A]_{\mu}\). Let \( \hat{\mathcal{B}} = \{[A]_{\mu} : A \in \mathcal{B}\} \) be the set of equivalence classes and define the following operations on \( \hat{\mathcal{B}} \): For all \( A, B \in \mathcal{B} \), let

\[ [A]_{\mu} = [A^c]_{\mu}, \ [A]_{\mu} \land [B]_{\mu} = [A \cap B]_{\mu}, \text{ and } [A]_{\mu} \lor [B]_{\mu} = [A \cup B]_{\mu}. \]

For all \((A_n)_{n \in \mathbb{N}} \in \mathcal{B}^\mathbb{N}\), let

\[ \bigwedge_{n=0}^{\infty} [A_n]_{\mu} = \left[ \bigcap_{n=0}^{\infty} A_n \right]_{\mu} \] and

\[ \bigvee_{n=0}^{\infty} [A_n]_{\mu} = \left[ \bigcup_{n=0}^{\infty} A_n \right]_{\mu}. \]

**4.2. Proposition.** As defined above, \( \land, \lor, \) and \( \land \) are well-defined.

**Proof.** Let \((X, \mathcal{B}, \mu)\) be a probability space and let \( \hat{\mathcal{B}} = \{[A]_{\mu} : A \in \mathcal{B}\} \). Let \([A]_{\mu} = [B]_{\mu} \in \hat{\mathcal{B}}. \) Since \( A \Delta B = A^c \Delta B^c, 0 = \mu(A \Delta B) = \mu(A^c \Delta B^c), \) so \([A]_{\mu} = [B]_{\mu}. \) Next, let \((A_n)_{n \in \mathbb{N}}, (B_n)_{n \in \mathbb{N}} \in \mathcal{B}^\mathbb{N}\) such that \([A_n]_{\mu} = [B_n]_{\mu} \) for all \( n \in \mathbb{N}. \) Then \( \mu((\bigcup_{i=0}^{\infty} A_i) \Delta (\bigcup_{j=0}^{\infty} B_j)) \)

\[ = \mu((\bigcup_{i=0}^{\infty} A_i) \setminus (\bigcup_{j=0}^{\infty} B_j)) \cup ((\bigcup_{j=0}^{\infty} B_j) \setminus (\bigcup_{i=0}^{\infty} A_i))) \]

\[ = \mu(\bigcup_{i=0}^{\infty} (A_i \setminus B_j)) \cup \bigcup_{j=0}^{\infty} (B_j \setminus A_i)) \]

\[ \leq \mu(\bigcup_{i=0}^{\infty} (A_i \setminus B_j)) + \mu(\bigcup_{j=0}^{\infty} (B_j \setminus A_i)) \]

\[ \leq \sum_{i=0}^{\infty} \mu(A_i \setminus B_i) + \sum_{j=0}^{\infty} \mu(B_j \setminus A_j) \]

\[ = 0, \]

so \( \bigcup_{i=0}^{\infty} A_i \sim_{\mu} \bigcup_{j=0}^{\infty} B_j. \) Thus \([\bigcup_{i=0}^{\infty} A_i]_{\mu} = [\bigcup_{j=0}^{\infty} B_j]_{\mu}\) and \( \lor \) is well-defined. By the above,

\([A_n^c]_{\mu} = [B_n^c]_{\mu}\) for all \( n \in \mathbb{N}, \) so \( \mu((\bigcup_{i=0}^{\infty} A_i^c) \Delta (\bigcup_{j=0}^{\infty} B_j^c)) = 0. \) Then \( \mu((\bigcap_{i=0}^{\infty} A_i) \Delta (\bigcap_{j=0}^{\infty} B_j)) =\)
\[
\mu \left( \bigcup_{i=0}^{\infty} A_i^c \right) \triangle \left( \bigcup_{j=0}^{\infty} B_j^c \right) = \mu \left( \bigcup_{i=0}^{\infty} A_i^c \right) \triangle \left( \bigcup_{j=0}^{\infty} B_j^c \right) = 0, \text{ so } \bigcap_{i=0}^{\infty} A_i \sim \mu \bigcap_{j=0}^{\infty} B_j. \text{ Thus } \bigcap_{j=0}^{\infty} B_j \mu \text{ and } \wedge \text{ is well-defined.} \]

4.3. Proposition. The set \( \hat{\mathcal{B}} \) with \( \emptyset = [\emptyset]_\mu, 1 = [X]_\mu \) and operations \( \wedge, \vee, \text{ and } \cdot^c \) is a Boolean \( \sigma \)-algebra.

Proof. Let \([A]_\mu, [B]_\mu, \text{ and } [C]_\mu \in \hat{\mathcal{B}}.

\begin{enumerate}
\item \([A]_\mu \vee [A]_\mu = [A \cup A]_\mu = [A]_\mu\)
\item \([A]_\mu \wedge [A]_\mu = [A \cap A]_\mu = [A]_\mu\)
\item \([A]_\mu \vee [B]_\mu = [A \cup B]_\mu = [B \cup A]_\mu = [B]_\mu \vee [A]_\mu\)
\item \([A]_\mu \wedge [B]_\mu = [A \cap B]_\mu = [B \cap A]_\mu = [B]_\mu \wedge [A]_\mu\)
\item \(([A]_\mu \vee [B]_\mu) \vee [B]_\mu = [A \cap B]_\mu \vee [B]_\mu = [(A \cap B) \cup B]_\mu = [B]_\mu\)
\item \(([A]_\mu \vee [B]_\mu) \wedge [B]_\mu = [A \cup B]_\mu \wedge [B]_\mu = [(A \cup B) \cap B]_\mu = [B]_\mu\)
\item \(([A]_\mu \vee [B]_\mu) \vee [C]_\mu = [A \cup B]_\mu \vee [C]_\mu = [(A \cup B) \cup C]_\mu = [A \cup (B \cup C)]_\mu = [A]_\mu \vee [(B \cup C)]_\mu\)
\item \(([A]_\mu \wedge [B]_\mu) \wedge [C]_\mu = [A \cap B]_\mu \wedge [C]_\mu = [(A \cap B) \cap C]_\mu = [A \cap (B \cap C)]_\mu = [A]_\mu \wedge ([B]_\mu \wedge [C]_\mu)\)
\item \([A]_\mu \vee ([B]_\mu \wedge [C]_\mu) = [A]_\mu \vee [B \cap C]_\mu = [A \cup (B \cap C)]_\mu = [(A \cup B) \cap (A \cup C)]_\mu = [A \cap B]_\mu \wedge [A \cup C]_\mu = ([A]_\mu \vee [B]_\mu) \wedge ([A]_\mu \vee [C]_\mu)\)
\item \([A]_\mu \wedge ([B]_\mu \vee [C]_\mu) = [A]_\mu \wedge [B \cup C]_\mu = [A \cap (B \cup C)]_\mu = [(A \cap B) \cup (A \cap C)]_\mu = [A \cap B]_\mu \vee [A \cap C]_\mu = ([A]_\mu \wedge [B]_\mu) \vee ([A]_\mu \wedge [C]_\mu)\)
\item \([A]_\mu \vee [A]^c = [A]_\mu \vee [A]^c = [A \cup A]^c = [X]_\mu\)
\item \([A]^c \wedge [A]_\mu = [A^c \cap A]_\mu = [\emptyset]_\mu\)
\item \([\emptyset]_\mu \neq [X]_\mu \text{ since } \mu((\emptyset \triangle X) = 1 > 0\]
Let $\langle [A_i]_\mu \rangle \in \mathcal{B}^N$. Let $B = \bigvee_{i=0}^{\infty} [A_i]_\mu$. For all $n \in \mathbb{N}$, $[A_n]_\mu \land B = [A_n]_\mu \land (\bigvee_{i=0}^{\infty} [A_i]_\mu) = [A_n]_\mu \land \left( \bigcup_{i=0}^{\infty} A_i \right)_\mu = \left[ (A_n \cap A_i) \mu \cup (A_n \cap A_i) \mu_n \right] = [A_n]_\mu$, so $[A_n]_\mu \leq B$.

Thus $B$ is an upper bound for $\langle [A_i]_\mu \rangle$. Now suppose $[C]_\mu = \hat{C} \in \mathcal{B}$ is another upper bound for $\langle [A_i]_\mu \rangle$. Then $B \land \hat{C} = \left( \bigvee_{i=0}^{\infty} [A_i]_\mu \right) \land [C]_\mu = \left[ \bigcup_{i=0}^{\infty} A_i \right]_\mu \land [C]_\mu = \left[ (\bigcup_{i=0}^{\infty} A_i) \cap C \right]_\mu = \left[ \bigcup_{i=0}^{\infty} (A_i \cap C) \right]_\mu = \bigvee_{i=0}^{\infty} (\left[ A_i \right]_\mu \land [C]_\mu) = \bigvee_{i=0}^{\infty} ([A_i]_\mu \land \hat{C})$. Since for all $i \in \mathbb{N}$, $\left[ A_i \right]_\mu \leq \hat{C}$, we have $B \land \hat{C} = \bigvee_{i=0}^{\infty} [A_i]_\mu = B$, so $B \leq \hat{C}$. Hence $B$ is the least upper bound of $\langle [A_i] \rangle$, so $\mathcal{B}$ is a Boolean $\sigma$-algebra.

Let $\hat{\mu} : \mathcal{B} \to [0,1]$ such that $\hat{\mu}(\{A\}_\mu) = \mu(A)$.

**4.4. Proposition.** $\hat{\mu}$ is well-defined and strictly positive.

**Proof.** Let $A, C \in \mathcal{B}$ such that $[A]_\mu = [C]_\mu$. By definition, $A \sim \mu C$, so $\mu(A \setminus C) + \mu(C \setminus A) = \mu(A \Delta C) = 0$. Then $\mu(A \setminus C) + \mu(C \setminus A) = \mu(A \cup (C \setminus A)) = \mu(A \cup C) = \mu(C \cup (A \setminus C)) = \mu(C) + \mu(A \setminus C) = \mu(C)$. Hence $\hat{\mu}(\{A\}_\mu) = \mu(A) = \mu(C) = \hat{\mu}(\{C\}_\mu)$, so $\hat{\mu}$ is well-defined. Suppose $\hat{\mu}(\{A\}_\mu) = 0$. Then $\mu(A) = 0$, so $\mu(A \Delta \emptyset) = 0$. Thus $A \sim \emptyset$, so $[A]_\mu = \emptyset = [\emptyset]_\mu = 0$.

**4.5. Proposition.** $\hat{\mu}$ is countably additive.

**Proof.** Let $(B_n)_{n \in \mathbb{N}} \in \mathcal{B}^N$ be pairwise disjoint, i.e. for all $m \neq n$, $B_m \land B_n = \emptyset$. For each $n \in \mathbb{N}$, choose $A_n \in \mathcal{B}$ such that $[A_n]_\mu = B_n$. Let $m, n \in \mathbb{N}$. Then $[\emptyset]_\mu = 0 = [A_m]_\mu \land [A_n]_\mu = [A_m \cap A_n]_\mu$, so $\emptyset \sim A_m \cap A_n$. Therefore, $0 = \mu((A_m \cap A_n) \setminus \emptyset) = \mu((A_m \setminus \emptyset) \cup (A_n \setminus \emptyset)) = \mu(A_m \cap A_n) + 0$, so $\mu(A_m \cap A_n) = 0$.

To use the countable additivity of $\mu$, we will construct a disjoint sequence $(D_n)_{n \in \mathbb{N}}$ in $\mathcal{B}$ such that $\mu(D_n) = \mu(A_n)$ for all $n \in \mathbb{N}$. Let $D_0 = A_0$ and $D_n = A_n \setminus \bigcup_{i=0}^{n-1} A_i$ for all
\( n > 0 \) and let \( n \in \mathbb{N} \). Then

\[
\mu(D_n) = \mu(A_n \setminus (\bigcup_{i=0}^{n-1} A_i))
\]

\[
= \mu(A_n \setminus (A_n \cap (\bigcup_{i=0}^{n-1} A_i)))
\]

\[
= \mu(A_n \setminus (\bigcup_{i=0}^{n-1} (A_n \cap A_i)))
\]

\[
= \mu(A_n) - \mu(\bigcup_{i=0}^{n-1} (A_n \cap A_i)) \quad \text{since} \quad \bigcup_{i=0}^{n-1} (A_n \cap A_i) \subseteq A_n
\]

\[
\leq \mu(A_n) - \sum_{i=0}^{n-1} \mu(A_n \cap A_i)
\]

\[
= \mu(A_n) \quad \text{since} \quad \mu(A_n \cap A_i) = 0 \text{ for all } i \in \mathbb{N},
\]

so

\[
\sum_{i=0}^{\infty} \hat{\mu}B_i = \sum_{i=0}^{\infty} \hat{\mu}[A_i]_{\mu}
\]

\[
= \sum_{i=0}^{\infty} \mu A_i
\]

\[
= \sum_{i=0}^{\infty} \mu D_i
\]

\[
= \mu(\bigcup_{i=0}^{\infty} D_i) \quad \text{by countable additivity of } \mu
\]

\[
= \mu(\bigcup_{i=0}^{\infty} A_i)
\]

\[
= \hat{\mu}(\bigcup_{i=0}^{\infty} [A_i]_{\mu})
\]

\[
= \hat{\mu}(\bigcup_{i \in \mathbb{N}} [A_i]_{\mu}) \quad \text{by definition}
\]

23
\[ = \hat{\mu}(\bigvee_{i \in \mathbb{N}} B_i) \]

Thus \(\hat{\mu}\) is countably additive. \qed

Since \(\hat{\mu}([\emptyset]) = 0\), \(\hat{\mu}([X]) = \mu(X) = 1\), and \(\hat{\mu}\) is countably additive, \(\hat{\mu}\) is a probability measure.

Let \(\hat{d} : \hat{\mathcal{B}} \rightarrow \mathbb{R}\) such that \(\hat{d}([A]_\mu, [B]_\mu) = \hat{\mu}([A]_\mu \triangle [B]_\mu)\)

4.6. **Proposition.** \(\hat{d}\) is a metric on \(\hat{\mathcal{B}}\).

**Proof.** Let \([A]_\mu, [B]_\mu, [C]_\mu \in \hat{\mathcal{B}}\).

(1) non-negativity: \(\hat{d}([A]_\mu, [B]_\mu) = \hat{\mu}([A]_\mu \triangle [B]_\mu) \geq 0\)

(2) coincidence: \(\hat{d}([A]_\mu, [B]_\mu) = 0 \iff \hat{\mu}([A]_\mu \triangle [B]_\mu) = \hat{\mu}([A \triangle B]_\mu) = 0 \iff \mu(A \triangle B) = 0 \iff A \sim_\mu B \iff [A]_\mu = [B]_\mu\)

(3) symmetry: \(\hat{d}([A]_\mu, [B]_\mu) = \hat{\mu}([A]_\mu \triangle [B]_\mu) = \hat{\mu}([A \triangle B]_\mu) = \hat{\mu}([B \triangle A]_\mu) = \hat{\mu}([B]_\mu \triangle [A]_\mu) = \hat{d}([B]_\mu, [A]_\mu)\)

(4) triangle inequality: \(\hat{d}([A]_\mu, [B]_\mu) + \hat{d}([B]_\mu, [C]_\mu) \geq \hat{d}([A]_\mu, [C]_\mu)\)

\[ = \hat{\mu}([A]_\mu \triangle [B]_\mu) + \hat{\mu}([B]_\mu \triangle [C]_\mu) \]

\[ = \mu(A \triangle B) + \mu(B \triangle C) \]

\[ = \mu(A \setminus B) + \mu(B \setminus A) + \mu(B \setminus C) + \mu(C \setminus B) \text{ since } (A \setminus B) \cap (B \setminus A) = \emptyset \]

\[ = (B \setminus C) \cap (C \setminus B) \]

\[ \geq \mu(A \setminus C) + \mu(C \setminus A) \]

since \(A \setminus C \subseteq (A \setminus B) \cup (B \setminus C)\) and \(C \setminus A \subseteq (C \setminus B) \cup (B \setminus A)\)

\[ = \mu(A \triangle C) \text{ since } (A \setminus C) \cap (C \setminus A) = \emptyset \]
\[= \hat{\mu}(\triangle [A]_{\mu}) \]
\[= \hat{d}(\triangle [A]_{\mu}, [C]_{\mu}). \]

4.7. Proposition. \((\hat{B}, \hat{d})\) is a complete metric space.

Proof. (323F in [9]) Let \(\langle a_n \rangle \in \hat{B}^\mathbb{N}\) be Cauchy with respect to \(\hat{d}\). For all \(k \in \mathbb{N}\) there exists a minimal \(N_k \in \mathbb{N}\) such that for all \(m, n \geq N_k\), \(\hat{d}(a_m, a_n) < 2^{-k}\). Note that \(\langle N_k \rangle_{k \in \mathbb{N}}\) is a non-decreasing sequence. For each \(k\), let \(c_k = a_{N_k}\). Then \(\sum_{k=0}^{\infty} \hat{d}(c_k, c_{k+1}) \leq \sum_{k=0}^{\infty} 2^{-k} < \infty\).

Let \(p_0 = \bigvee_{n \in \mathbb{N}} \bigwedge_{m \geq n} c_m\) and let \(p_1 = \bigwedge_{n \in \mathbb{N}} \bigvee_{m \geq n} c_m\). Note that since \(\hat{B}\) is \(\sigma\)-complete, \(p_0, p_1 \in \hat{B}\).

First, I will show that \(p_0 = p_1\). Let \(\alpha_n = \hat{\mu}(c_n \triangle c_{n+1})\) and \(\beta_n = \sum_{k=n}^{\infty} \alpha_k\) for all \(n \in \mathbb{N}\). Since \(\hat{d}(c_n, c_{n+1}) = \hat{\mu}(c_n \triangle c_{n+1}) = \alpha_n\) for all \(n\), \(\sum_{n=0}^{\infty} \alpha_n = \sum_{k=0}^{\infty} \hat{d}(c_k, c_{k+1}) < \infty\), so \(\lim_{n \to \infty} \beta_n = 0\). Let \(b_n = \bigvee_{n \leq k} c_k \triangle c_{k+1}\). Then for all \(n\),

\[\hat{\mu}(b_n) = \hat{\mu}(\bigvee_{m \geq n} c_m \triangle c_{m+1}) \leq \sum_{m=n}^{\infty} \hat{\mu}(c_m \triangle c_{m+1}) = \beta_n\]

by countable additivity of \(\hat{\mu}\).

Fix \(n \in \mathbb{N}\) and let \(n \leq m\). I will show that \(c_m \triangle c_n \leq \bigvee_{n \leq k < m} c_k \triangle c_{k+1}\). If \(m = n\), then

\[c_m \triangle c_n = 0 = \bigvee_{n \leq k < m} c_k \triangle c_{k+1} = \bigvee_{n \leq k < m} c_k \triangle c_{k+1}\]

If \(m = n + 1\), then \(c_m \triangle c_n = c_{n+1} \triangle c_n = \bigvee_{k=n} c_k \triangle c_{k+1}\). For \(m \geq n + 2\), \(\bigvee_{n \leq k < m} c_k \triangle c_{k+1}\)

\[= \left( \bigvee_{n \leq k < m-2} c_k \triangle c_{k+1} \right) \cup \left( (c_{m-2} \triangle c_{m-1}) \cup (c_{m-1} \triangle c_m) \right) \]

\[\geq \left( \bigvee_{n \leq k < m-2} c_k \triangle c_{k+1} \right) \cup (c_{m-2} \triangle c_m) \text{ by 4.1.5}\]
Thus, \( c_m \triangle c_n \leq \bigvee_{n < k < m} c_k \triangle c_{k+1} \leq \bigvee_{n \leq k < m} c_k \triangle c_{k+1} = b_n \). Since \( c_m \triangle c_n \leq b_n \), we have
\[
\begin{aligned}
(c_n \setminus b_n) \setminus c_m &= (c_n \setminus b_n \setminus c_m) = (c_n \setminus c_m) \setminus c_m \setminus c_n \leq (c_n \setminus c_m) \setminus b_n \leq b_n \setminus b_n = 0,
\end{aligned}
\]
so \((c_n \setminus b_n) \setminus c_m = 0\). By 4.1.6, \( c_n \setminus b_n \leq c_m \).

We can also use \( c_m \triangle c_n \leq b_n \) to show that \( c_m \leq c_n \lor b_n \). Since \( c_m \setminus c_n \leq b_n \), \((c_m \setminus c_n) \land b_n = c_m \setminus c_n\), so \((c_m \setminus c_n) \land b_n \lor c_n = (c_m \setminus c_n) \lor c_n \). Expanding both sides, we get
\[
\begin{aligned}
(c_m \land c_n \land b_n) \lor c_n &= (c_m \land c_n) \lor c_n, \quad \text{so} \quad (c_m \lor c_n) \land (b_n \lor c_n) = (c_m \lor c_n) \land 1 \land (b_n \lor c_n) = (c_m \lor c_n) \land (b_n \lor c_n) = (c_m \lor c_n) \land 1 = c_m \lor c_n. \quad \text{Thus} \quad c_m \leq c_m \lor c_n \leq c_n \lor b_n.
\end{aligned}
\]

Putting this together, we get
\[
c_n \setminus b_n \leq c_m \leq c_n \lor b_n.
\]

Since this holds for every \( m \geq n \),
\[
c_n \setminus b_n \leq \bigwedge_{m \geq k} c_m \leq \bigvee_{m \geq k} c_m \leq c_n \lor b_n
\]
for all \( k \geq n \). Since this holds for every \( k \geq n \), \( c_n \setminus b_n \leq \bigvee_{k \in \mathbb{N}} \bigwedge_{m \geq k} c_m \leq \bigwedge_{k \in \mathbb{N}} \bigvee_{m \geq k} c_m \leq c_n \lor b_n \). In other words, since \( p_0 = \bigvee_{k \in \mathbb{N}} \bigwedge_{m \geq k} c_m \) and \( p_1 = \bigwedge_{k \in \mathbb{N}} \bigvee_{m \geq k} c_m \), we have
\[
\begin{aligned}
c_n \setminus b_n &\leq p_0 \leq p_1 \leq c_n \lor b_n.
\end{aligned}
\]

I will use this to show that \( c_n \triangle p_i \leq b_n \) for \( i = 0, 1 \). Let \( i \in \{0, 1\} \). Since \( c_n \setminus b_n \leq p_i \), we have \((c_n \setminus p_i) \setminus b_n = (c_n \setminus b_n) \setminus p_i \leq p_i \setminus p_i = 0\), so \((c_n \setminus p_i) \setminus b_n = 0\). By 4.1.6,
\( c_n \setminus p_i \leq b_n \). Since \( p_i \leq c_n \lor b_n \), \( p_i \land (c_n \lor b_n) = p_i \), so \( (p_i \setminus c_n) \setminus b_n = p_i \land c_n^c \land b_n^c = p_i \land (c_n \lor b_n)^c = (p_i \land (c_n \lor b_n)) \land (c_n \lor b_n)^c = 0 \). By 4.1.6, \( p_i \setminus c_n \leq b_n \). Thus \( c_n \triangle p_i = (c_n \setminus p_i) \lor (p_i \setminus c_n) \leq b_n \).

Another application of 4.1.6 gives us \( p_1 \setminus p_0 \leq b_n \):

\[
(p_1 \setminus p_0) \setminus b_n = p_1 \land p_0^c \land b_n^c
= (p_1 \land (c_n \lor b_n)) \land p_0^c \land b_n^c \text{ since } p_1 \leq c_n \lor b_n
= p_1 \land p_0^c \land (b_n^c \lor (c_n \lor b_n))
= p_1 \land p_0^c \land ((b_n^c \lor c_n) \lor (b_n^c \lor b_n))
= p_1 \land p_0^c \land ((b_n^c \lor c_n) \lor 0)
= p_1 \land p_0^c \land (c_n \setminus b_n)
= p_1 \land p_0^c \land ((c_n \setminus b_n) \land p_0) \text{ since } c_n \setminus b_n \leq p_0
= p_1 \land (c_n \setminus b_n) \land (p_0^c \land p_0)
= p_1 \land (c_n \setminus b_n) \land 0
= 0.
\]

Since \( c_n \triangle p_i \leq b_n \) holds for every \( n \),

\[
\lim_{n \to \infty} \hat{\mu}(c_n \triangle p_i) \leq \lim_{n \to \infty} \hat{\mu}(b_n) \leq \lim_{n \to \infty} \beta_n = 0
\]

for \( i = 0, 1 \). Furthermore, since \( p_0 \leq p_1 \), \( \hat{\mu}(p_1 \triangle p_0) = \hat{\mu}(p_1 \setminus p_0) \leq \lim_{n \to \infty} \hat{\mu}(b_n) = 0 \).

In fact, this implies that \( p_0 = p_1 \). To see this, let \( P_0, P_1 \in \mathcal{B} \) such that \( p_0 = [P_0], p_1 = [P_1] \). Then \( 0 = \hat{\mu}(p_1 \triangle p_0) = \hat{\mu}([P_1] \triangle [P_0]) = \hat{\mu}([P_1 \triangle P_0]) = \mu(P_1 \triangle P_0) \), so \( p_0 = [P_0] = [P_1] = p_1 \).
Now let $p = p_0 = p_1$. I will show that $\lim_{n \to \infty} a_n = \lim_{n \to \infty} c_n = p$. Since $\lim_{n \to \infty} \hat{d}(c_n, p) = \lim_{n \to \infty} \hat{\mu}(c_n \triangle p) = 0$, $\langle c_n \rangle_{n \in \mathbb{N}}$ converges to $p$ with respect to $\hat{d}$. Let $\epsilon > 0$. Since $\langle c_n \rangle_{n \in \mathbb{N}}$ converges to $p$, there exists $M_1 \in \mathbb{N}$ such that for all $k > M_1$, $\hat{d}(c_k, p) < \epsilon/2$. Since $\langle a_n \rangle_{n \in \mathbb{N}}$ is Cauchy, there exists $M_2 \in \mathbb{N}$ such that for all $m, n > M_2$, $\hat{d}(a_m, a_n) < \epsilon/2$. Let $M = \max(M_1, M_2)$. Then $k > M$ implies $\hat{d}(a_k, p) \leq \hat{d}(a_k, a_{N_k}) + \hat{d}(a_{N_k}, p) = \hat{d}(a_k, a_{N_k}) + \hat{d}(c_k, p) < \epsilon/2 + \epsilon/2 = \epsilon$, so the original sequence $\langle a_n \rangle_{n \in \mathbb{N}}$ converges to $p \in \hat{B}$. Since $\langle a_n \rangle_{n \in \mathbb{N}}$ was an arbitrary Cauchy sequence, $(\hat{B}, \hat{d})$ is a complete metric space. □

Since $(\hat{B}, \hat{d})$ is a complete metric space, $\mathcal{M} = (\hat{B}, 0, 1, \wedge, \vee, \hat{\mu})$ as constructed above is a metric structure. We call such an $\mathcal{M}$ a probability algebra, or, in the context of its being a metric structure, a probability structure.

### 4.8. Here are some examples of probability structures.

1. Let $(X, \mathcal{B}, \mu)$ be a probability space where $X = \{a_1, a_2, \ldots, a_{10}\}$, $\mathcal{B} = \mathcal{P}(X)$, and $\mu(a_n) = 1/10$ for all $n = 1, \ldots, 10$. Since there are no nontrivial events of measure zero, $\mu = \hat{\mu}$. The distance between two events is the probability that exactly one of them happens.

2. Let $(X, \mathcal{B}, \mu)$ be a probability space where $X$ is the unit interval $[0, 1]$, $\mathcal{B}$ is the $\sigma$-algebra of Borel sets, and $\mu$ is the Lebesgue measure on $[0, 1]$. $\hat{B}$ identifies events whose symmetric difference is a null set.

3. Let $X = \{0, 1\}^\mathbb{N}$, the set of infinite sequences of 0s and 1s. $\mathcal{B}$ is the $\sigma$-algebra generated by sets of the form $\{(x_1, x_2, \ldots) \in X : (x_1, \ldots, x_n) = (a_1, \ldots, a_n)\}$. If we think of $X$ as the set of countably many tosses of a fair coin, then a generator
of $\mathcal{B}$ is the event that the first $n$ tosses are a fixed sequence $(a_1, ..., a_n)$. Then 
\[ \mu((x_1, x_2, ...) \in X : (x_1, ..., x_n) = (a_1, ..., a_n)) = 2^{-n}. \] Two distinct generator sets have symmetric difference $> 0$, and thus remain distinct in $\hat{\mathcal{B}}$.

**Definition.** The signature associated with probability structures is $L^{Pr} = (\mathcal{R}, \mathcal{F}, \eta, \mathcal{G})$, where

1. $\mathcal{R} = \{\hat{\mu}, \hat{d}\}$
2. $\mathcal{F} = \{0, 1, \land, \lor\}$
3. $0 = \eta(0) = \eta(1), 1 = \eta(\hat{\mu}) = \eta(\land), 2 = \eta(\hat{d}) = \eta(\lor) = \eta(\land)$
4. $\mathcal{G} = \{\delta_\epsilon, \delta_\mu, \delta_\lor, \delta_\land, \delta_\delta\}$ where $\delta_\epsilon(\epsilon) = \delta_\mu(\epsilon) = \epsilon$ and $\delta_\lor(\epsilon) = \delta_\land(\epsilon) = \delta_\delta(\epsilon) = \epsilon/2$.

**4.9. Proposition.** As defined above, $\delta_\epsilon, \delta_\mu, \delta_\lor, \delta_\land, \delta_\delta$ are indeed moduli of uniform continuity.

**Proof.** $\delta_\epsilon$: Let $\epsilon \in (0, 1]$ and $[A]_\mu, [B]_\mu \in \hat{\mathcal{B}}$. Suppose $d([A]_\mu, [B]_\mu) < \delta_\epsilon(\epsilon) = \epsilon$. Then 
$\quad d([A]_\mu^c, [B]_\mu^c) = d([A^c]_\mu, [B^c]_\mu) = \mu(A \Delta B^c) = \mu(A \land B) = d([A]_\mu, [B]_\mu) < \epsilon.$

$\delta_\mu$: Let $\epsilon \in (0, 1]$ and $[A]_\mu, [B]_\mu \in \hat{\mathcal{B}}$. Suppose $d([A]_\mu, [B]_\mu) < \delta_\mu(\epsilon) = \epsilon$. Since 
$\quad 2\hat{\mu}(a \land b) \leq 2\hat{\mu}(a)$, we have $-\hat{\mu}(a) + 2\hat{\mu}(a \land b) \leq \hat{\mu}(a)$, so $-\hat{\mu}(a \land b) = -\hat{\mu}(a) - \hat{\mu}(b) + 2\hat{\mu}(a \land b) \leq \hat{\mu}(a) - \hat{\mu}(b)$. Since $2\hat{\mu}(a \land b) \leq 2\hat{\mu}(b)$, we have $-\hat{\mu}(b) \leq \hat{\mu}(b) - 2\hat{\mu}(a \land b)$, so $\hat{\mu}(a) - \hat{\mu}(b) \leq \hat{\mu}(a) + \hat{\mu}(b) - 2\hat{\mu}(a \land b) = \hat{\mu}(a \land b)$. Putting this together, we get 
$\quad d(\hat{\mu}(a), \hat{\mu}(b)) = |\hat{\mu}(a) - \hat{\mu}(b)| \leq \hat{\mu}(a \land b) < \epsilon.$

$\delta_\lor$: Let $\epsilon \in (0, 1]$ and $([A_1]_\mu, [A_2]_\mu), ([B_1]_\mu, [B_2]_\mu) \in \hat{\mathcal{B}} \times \hat{\mathcal{B}}$. Suppose 
$\quad d_{\max}(([A_1]_\mu, [A_2]_\mu), ([B_1]_\mu, [B_2]_\mu)) < \delta_\lor(\epsilon)$

where $\delta_\lor(\epsilon) = \epsilon/2$. Then for $i = 1, 2$, $\epsilon/2 > d([A_i]_\mu, [B_i]_\mu) = \mu(A_i \Delta B_i) = \mu(A_i \setminus B_i) + \mu(B_i \setminus A_i)$. Since $(A_1 \cup A_2) \setminus (B_1 \cup B_2) = (A_1 \setminus (B_1 \cup B_2)) \cup (A_2 \setminus (B_1 \cup B_2)) \subseteq$
(A_1 \setminus B_1) \cup (A_2 \setminus B_2) and (B_1 \cup B_2) \setminus (A_1 \cup A_2) = (B_1 \setminus (A_1 \cup A_2)) \cup (B_2 \setminus (A_1 \cup A_2)) \subseteq (B_1 \setminus A_1) \cup (B_2 \setminus A_2), we have \(d([A_1]_\mu \vee [A_2]_\mu, [B_1]_\mu \vee [B_2]_\mu)\)

\[
= d([A_1 \cup A_2]_\mu, [B_1 \cup B_2]_\mu)
\]

= \(\mu((A_1 \cup A_2) \triangle (B_1 \cup B_2))\)

= \(\mu((A_1 \cup A_2) \setminus (B_1 \cup B_2)) + \mu((B_1 \cup B_2) \setminus (A_1 \cup A_2))\)

\(\leq \mu((A_1 \setminus B_1) \cup (A_2 \setminus B_2)) + \mu((B_1 \setminus A_1) \cup (B_2 \setminus A_2))\)

\(\leq \mu(A_1 \setminus B_1) + \mu(A_2 \setminus B_2) + \mu(B_1 \setminus A_1) + \mu(B_2 \setminus A_2)\)

= \((\mu(A_1 \setminus B_1) + \mu(B_1 \setminus A_1)) + (\mu(A_2 \setminus B_2) + \mu(B_2 \setminus A_2))\)

= \(\mu(A_1 \triangle B_1) + \mu(A_2 \triangle B_2)\)

= \(d([A_1]_\mu, [B_1]_\mu) + d([A_2]_\mu, [B_2]_\mu)\)

\(\leq 2\max\{d([A_1]_\mu, [B_1]_\mu), d([A_2]_\mu, [B_2]_\mu)\}\)

< \(2(\epsilon/2)\)

= \(\epsilon.\)

\(\delta_\wedge: \) Let \(\epsilon \in (0, 1]\) and \(([A_1]_\mu, [A_2]_\mu), ([B_1]_\mu, [B_2]_\mu) \in \hat{\mathcal{B}} \times \hat{\mathcal{B}}.\) Suppose

\[
d_{\text{max}}(([A_1]_\mu, [A_2]_\mu), ([B_1]_\mu, [B_2]_\mu)) < \delta_\wedge(\epsilon)
\]

where \(\delta_\wedge(\epsilon) = \epsilon/2.\) Since \((A_1 \cap A_2) \setminus (B_1 \cap B_2) = ((A_1 \cap A_2) \setminus B_1) \cup ((A_1 \cap A_2) \setminus B_2) \subseteq (A_1 \setminus B_1) \cup (A_2 \setminus B_2)\) and \((B_1 \cap B_2) \setminus (A_1 \cap A_2) = ((B_1 \cap B_2) \setminus A_1) \cup ((B_1 \cap B_2) \setminus A_2) \subseteq (B_1 \setminus A_1) \cup (B_2 \setminus A_2),\) we have \(d([A_1]_\mu \land [A_2]_\mu, [B_1]_\mu \land [B_2]_\mu)\)

\[
= d([A_1 \cap A_2]_\mu, [B_1 \cap B_2]_\mu)
\]

30
\[ \mu((A_1 \cap A_2) \triangle (B_1 \cap B_2)) = \mu((A_1 \cap A_2) \setminus (B_1 \cap B_2)) + \mu((B_1 \cap B_2) \setminus (A_1 \cap A_2)) \leq \mu((A_1 \setminus B_1) \cup (A_2 \setminus B_2)) + \mu((B_1 \setminus A_1) \cup (B_2 \setminus A_2)) \leq \mu(A_1 \setminus B_1) + \mu(A_2 \setminus B_2) + \mu(B_1 \setminus A_1) + \mu(B_2 \setminus A_2) \]

\[ = (\mu(A_1 \setminus B_1) + \mu(B_1 \setminus A_1)) + (\mu(A_2 \setminus B_2) + \mu(B_2 \setminus A_2)) = \mu(A_1 \triangle B_1) + \mu(A_2 \triangle B_2) \]

\[ = d([A_1]_\mu, [B_1]_\mu) + d([A_2]_\mu, [B_2]_\mu) \leq 2\max\{d([A_1]_\mu, [B_1]_\mu), d([A_2]_\mu, [B_2]_\mu)\} < 2(\epsilon/2) = \epsilon. \]

\[ \delta_d^\epsilon: \text{ Let } \epsilon \in (0, 1] \text{ and } ([A_1]_\mu, [A_2]_\mu), ([B_1]_\mu, [B_2]_\mu) \in \hat{\mathcal{B}} \times \hat{\mathcal{B}}. \text{ Suppose} \]

\[ d_{\max}(([A_1]_\mu, [A_2]_\mu), ([B_1]_\mu, [B_2]_\mu)) < \delta_d^\epsilon(\epsilon) \]

where \( \delta_d^\epsilon(\epsilon) = \epsilon/2 \). By the triangle inequality,

\[ \hat{d}(([A_1]_\mu, [A_2]_\mu)) - \hat{d}(([B_1]_\mu, [B_2]_\mu)) \leq \hat{d}(([A_1]_\mu, [B_1]_\mu)) + \hat{d}(([B_1]_\mu, [B_2]_\mu)) + \hat{d}(([B_2]_\mu, [A_2]_\mu)) - \hat{d}(([B_1]_\mu, [B_2]_\mu)) = \hat{d}(([A_1]_\mu, [B_1]_\mu)) + \hat{d}(([B_2]_\mu, [A_2]_\mu)) < \epsilon/2 + \epsilon/2 = \epsilon. \]
since \( \max\{d([A_1]_\mu, [B_1]_\mu), d([A_2]_\mu, [B_2]_\mu)\} < \epsilon/2 \). Similarly,

\[
\hat{d}(([A_1]_\mu, [A_2]_\mu)) - \hat{d}(([B_1]_\mu, [B_2]_\mu)) \\
\geq \hat{d}(([A_1]_\mu, [A_2]_\mu)) - (\hat{d}(([B_1]_\mu, [A_1]_\mu)) + \hat{d}(([A_1]_\mu, [A_2]_\mu)) + \hat{d}(([A_2]_\mu, [B_2]_\mu))) \\
= -(\hat{d}(([B_1]_\mu, [A_1]_\mu)) + \hat{d}(([A_2]_\mu, [B_2]_\mu))) \\
> -(\epsilon/2 + \epsilon/2) \\
= -\epsilon.
\]

Putting this together, we get

\[
d(\hat{d}(([A_1]_\mu, [A_2]_\mu)), \hat{d}(([B_1]_\mu, [B_2]_\mu))) = |\hat{d}(([A_1]_\mu, [A_2]_\mu)) - \hat{d}(([B_1]_\mu, [B_2]_\mu))| < \epsilon.
\]

\[\square\]

**Notation.** A Boolean ring is a ring \((A, +, \cdot)\) where \(a^2 = a\) for all \(a \in A\). We can represent any Boolean algebra \(A = (A, 0, 1, \wedge, \vee, c)\) as a Boolean ring \((A, +, \cdot)\) by taking \(a + b := a \triangle b\) and \(ab := a \wedge b\). Conversely, given a Boolean ring \((A, +, \cdot)\), we can view it as a Boolean algebra by taking \(a \vee b := a + b + ab, a \wedge b := ab,\) and \(a^c := 1 + a\). Depending on the context, one representation may be preferred over the other.

**Definition.** (311F in [9]) Given a Boolean ring \(B\), the Stone space of \(B\) is the set \(Z\) of nonzero ring homomorphisms from \(B\) to \(\mathbb{Z}_2\), i.e. \(z \in Z \iff \forall a, b \in B \ z(a + b) = z(a) + z(b), z(ab) = z(a)z(b), z(1) = 1_{\mathbb{Z}_2}\). For each \(a \in B\), let \(\hat{a} = \{z \in Z : z(a) = 1\}\).

The Stone representation of \(B\) is the canonical map \(a \mapsto \hat{a} : B \rightarrow \mathcal{P}(Z)\).

**4.10. Lemma.** (a special case of 311D in [9]) Let \(B\) be a Boolean ring and let \(a \in B\) be nonzero. Then there exists \(z \in Z\) such that \(z(a) = 1\).

**Proof.** Let \(\mathcal{I} = \{J : J\ is\ an\ ideal\ of\ B\ and\ a \notin J\}\). First I will use Zorn’s lemma on \((\mathcal{I}, \subseteq)\) to show that \(\mathcal{I}\) has a maximal element. Note that \(\mathcal{I}\) is nonempty because \(\{0\} \in \mathcal{I}\). Let \(\mathcal{J}\)
be a nonempty totally ordered subset of $\mathcal{I}$ and let $J_0 = \bigcup \mathcal{J}$. Let $m, n \in J_0$. Then there exist $J_1, J_2 \in \mathcal{J}$ such that $m \in J_1$ and $n \in J_2$. Since $\mathcal{J}$ is totally ordered, $J_1 \subseteq J_2$ or $J_2 \subseteq J_1$. Without loss of generality, suppose $J_1 \subseteq J_2$, so $m, n \in J_2$. Let $r \in \mathcal{B}$. Since $J_2$ is an ideal, $0, m + n, rm \in J_2 \subseteq J_0$, so $J_0$ is an ideal. Since $a \notin J_0$, $0 \in \mathcal{I}$, so $J_0$ is an upper bound for $\mathcal{J}$. Since $\mathcal{J}$ was arbitrary, by Zorn’s lemma, $\mathcal{I}$ has a maximal element, which we will denote by $K$.

Next, I will show that for all $m, n \in \mathcal{B} \setminus K$, $m + n \in K$. For all $b \in \mathcal{B}$, let $K_b = \{c \in \mathcal{B} : bc \in K\}$. Let $b \in \mathcal{B}$. Since $K$ is an ideal, for all $n \in K$, we have $nb \in K$, so $n \in K_b$. Thus $K \subseteq K_b$. If $n, n' \in K_b$ and $c \in \mathcal{B}$, then $b(n + n') = bn + bn' \in K$ and $b(nc) \in K$, so $n + n', nc \in K_b$. Since $0 \in K \subseteq K_b$, $K_b$ is an ideal. If $a \notin K_b$, then $K_b \in \mathcal{I}$. Since $K$ is maximal in $\mathcal{I}$ and $K \subseteq K_b$, we have $K = K_b$. In particular, $a \notin K_a$, since $a^2 = a \notin K$, so $K_a = K$. Now let $b, c \in \mathcal{B} \setminus K$. Since $b \notin K = K_a$, we have $ba = ab \notin K$, so $a \notin K_b$. Hence $K_b = K$. Since $c \notin K = K_b$, we have $bc \notin K$. Now $bc(b + c) = b^2c + bc^2 = bc + bc = 0 \in K$, so $b + c \in K_{bc}$. Since $bc \notin K$ and for all $e \notin K$, $K = K_e$, we have $K = K_{bc}$, so $b + c \in K$.

Now I will construct a ring homomorphism $z : \mathcal{B} \to \mathbb{Z}_2$ such that $z(a) = 1$. For all $d \in \mathcal{B}$, let

$$z(d) := \begin{cases} 0 & \text{if } d \in K \\ 1 & \text{if } d \in \mathcal{B} \setminus K. \end{cases}$$

Let $b, c \in \mathcal{B}$. If $b, c \in K$, then $b + c, bc \in K$, so $z(b + c) = 0 = 0 + 0 = z(b) + z(c)$ and $z(bc) = 0 = 0 \cdot 0 = z(b)z(c)$. If $b \in K$ and $c \in \mathcal{B} \setminus K$, then $b + c \notin K$, since otherwise $c = b + (b + c) \in K$, so $z(b + c) = 1 = 0 + 1 = z(b) + z(c)$. Since $bc \in K$, $z(bc) = 0 = 0 \cdot 1 = z(b)z(c)$. If $b \in \mathcal{B} \setminus K$ and $c \in K$, then $z(b + c) = 1 = 1 + 0 = z(b) + z(c)$.
and \( bc \in K \), \( z(bc) = 0 = 1 \cdot 0 = z(b)z(c) \). Finally, if \( b, c \in \mathcal{B} \setminus K \), then as shown in the previous paragraph, \( b + c \in K \) and \( bc \notin K \), so \( z(b + c) = 0 = 1 + 1 = z(b) + z(c) \) and \( z(bc) = 1 = 1 \cdot 1 = z(b)z(c) \). Hence \( z \) is a ring homomorphism and, since \( a \notin K \), \( z(a) = 1 \).

\[ \square \]

4.11. **Proposition.** (311E in [9]) The Stone representation of a Boolean ring \( \mathcal{B} \) is an injective ring homomorphism \( a \mapsto \hat{a} : (\mathcal{B}, +, \cdot) \to (\mathcal{P}(Z), \triangle, \cap) \).

**Proof.** Let \( \mathcal{B} \) be a Boolean algebra and let \( a, b \in \mathcal{B} \). Then \( \hat{a} + \hat{b} = \{ z : z(a + b) = 1 \} = \{ z : z(a) + z(b) = 1 \} = \{ z : z(a) = 0 \text{ and } z(b) = 1, \text{ or } z(a) = 1 \text{ and } z(b) = 0 \} = (\hat{a} \cap (\hat{b})^c) \cup (\hat{b} \cap (\hat{a})^c) = \hat{a} \triangle \hat{b} \) and \( \hat{a} \hat{b} = \{ z : z(ab) = 1 \} = \{ z : z(a)z(b) = 1 \} = \{ z : z(a) = 1 \text{ and } z(b) = 1 \} = \hat{a} \cap \hat{b} \), so \( a \mapsto \hat{a} \) is a ring homomorphism.

To see that this ring homomorphism is injective, let \( a \in \mathcal{B} \) be nonzero. By 4.10, there exists \( z \in Z \) such that \( z(a) = 1 \). Thus \( z \in \hat{a} \), so \( \hat{a} \neq \emptyset \). Hence the kernel of this homomorphism is \{0\}, so \( a \mapsto \hat{a} \) is injective.

\[ \square \]

By 4.11, any Boolean ring \( \mathcal{B} \) is isomorphic to its image \( \{ \hat{a} : a \in \mathcal{B} \} \) under the Stone representation. In particular, \( \triangle \) commutes with Boolean operations: for all \( a, b \in \mathcal{B} \),

1. \( \hat{a} \triangle b = a \triangle b \triangle (a \wedge b) = a + b + ab = \hat{a} \triangle b + ab = \hat{a} \triangle \hat{b} \triangle \hat{a} \triangle b = \hat{a} \cup \hat{b} \)
2. \( \hat{a} \wedge \hat{b} = \hat{a} \hat{b} = \hat{a} \triangle \hat{b} \)
3. \( \hat{a} \setminus \hat{b} = \hat{a} \triangle \hat{b} = a(1 + \hat{b}) = a + ab = \hat{a} \triangle \hat{b} = \hat{a} \triangle (\hat{a} \cap \hat{b}) = (\hat{a} \setminus (\hat{a} \cap \hat{b})) \cup ((\hat{a} \setminus \hat{b}) \setminus \hat{a}) = \hat{a} \setminus (\hat{a} \cap \hat{b}) = \hat{a} \setminus \hat{b} \)
4. \( \hat{a} \triangle b = a + b = \hat{a} \triangle \hat{b} \).
Also note that $a \mapsto \hat{a} : B \to \mathcal{P}(Z)$ respects the order relation on $B$, i.e. for all $a, b \in B$, $a \leq b \iff \hat{a} \subseteq \hat{b}$. Finally, we have $\hat{1} = Z$: Let $z \in Z$. By definition, $z$ is nonzero ring homomorphism, so there exists $a \in B$ such that $z(a) = 1$. Then $1 = z(a) = z(a \land 1) = z(a)z(1) = 1z(1) = z(1)$, so $z \in \hat{1} = Z$.

4.12. **Lemma.** Let $B$ be a Boolean algebra and let $Z$ be its Stone space. Let $\mathcal{T}$ be

$$\{G \subseteq Z : \forall z \in G \exists a \in B \ z \in \hat{a} \subseteq G\}.$$  

Then $\mathcal{T}$ is a topology on $Z$ under which $Z$ is a Hausdorff space.

**Proof.** First, we show that $\mathcal{T}$ is a topology. Let $z \in Z$. Since $z \neq 0$, there exists $a \in B$ such that $z(a) = 1$, so $z \in \hat{a}$. Thus $Z \subseteq \bigcup \mathcal{E}$. Now let $\hat{a}, \hat{b} \in \mathcal{E}$ and $z \in \hat{a} \cap \hat{b}$. By 311G, $\hat{a} \cap \hat{b} = a \land b \in \mathcal{E}$, so in particular, $z \in a \land b \subseteq a \cap b$. Thus $\mathcal{E}$ is a base for some topology on $Z$. Since $\mathcal{T}$ is defined to be the union of elements of $\mathcal{E}$, $\mathcal{T}$ is the topology generated by $\mathcal{E}$.

To see that $(Z, \mathcal{T})$ is Hausdorff, let $w, z \in Z$ be distinct. Then there exists $a \in B$ such that $z(a) \neq w(a)$, say $z(a) = 1$ and $w(a) = 0$. Since $w$ is nonzero, there exists $b \in B$ such that $w(b) = 1$. Then $w(b \setminus a) = w(b \land a^c) = w(b(1 + a)) = w(b + ba) = w(b) + w(b)w(a) = 1$, so $w \in b \setminus a \in \mathcal{T}$. Since $z(a) = 1$, $z \in \hat{a} \in \mathcal{T}$. Also, $(b \setminus a) \cap \hat{a} = (b \setminus a) \land a = b \land (a^c \land a) = 0 = \hat{1} \setminus \hat{1} = \hat{1} \setminus \hat{1} = \emptyset$, so $b \setminus a$ and $\hat{a}$ are disjoint neighborhoods of $w$ and $z$, respectively. Hence $(Z, \mathcal{T})$ is Hausdorff. \[\square\]

4.13. *(311I in [9])* The Stone space $Z$ of a Boolean algebra $B$ is compact. $\mathcal{E} = \{\hat{a} : a \in B\}$ is precisely the set of compact and open subsets of $Z$.

**Definition.** Let $X$ be a set and let $\mathcal{B}$ be a $\sigma$-algebra of subsets of $X$. $A \subseteq \mathcal{B}$ is a $\sigma$-ideal if it satisfies the following properties:
(1) $\emptyset \in A$

(2) For all $A \in A, B \in B$, if $B \subseteq A$, then $B \in A$

(3) If $\langle A_n \rangle_{n \in \mathbb{N}} \in A^\mathbb{N}$, then $\bigcup_{n \in \mathbb{N}} A_n \in A$

**Definition.** Let $X$ be a topological space. For $A \subseteq X$, denote the closure of $A$ by $\overline{A}$ and the interior of $A$ by $\text{int}A$. $A \subseteq X$ is nowhere dense if $\text{int}A = \emptyset$. $A \subseteq X$ is meager if there exist countably many $B_n \subseteq X$ such that $A = \bigcup_{n \in \mathbb{N}} B_n$ and each $B_n$ is nowhere dense.

**Definition.** (312F in [9]) Given Boolean algebras $A$ and $B$, a function $\pi : A \rightarrow B$ is a Boolean homomorphism if $\pi(a \Delta b) = \pi(a) \Delta \pi(b)$ and $\pi(a \cap b) = \pi(a) \cap \pi(b)$ for all $a, b \in A$, and $\pi(1_A) = 1_B$.

4.14. Lemma. (313Ca in [9]) Let $A$ be a $\sigma$-complete Boolean algebra and let $Z$ be its Stone space. Let $\langle a_n \rangle$ be a sequence in $A$ and let $\hat{a} = \sup_{n \in \mathbb{N}} a_n \in A$. Then $\bigcup_{n \in \mathbb{N}} \hat{a}_n = \hat{a}$.

**Proof.** For all $n$, $a_n \leq a$, so $\hat{a}_n \subseteq \hat{a}$ by 4.11. Then $\hat{a} \supseteq \bigcup_{n \in \mathbb{N}} \hat{a}_n = \bigcup_{n \in \mathbb{N}} \hat{a}_n$, since $\hat{a}$ is closed. To show that $\hat{a} \subseteq \bigcup_{n \in \mathbb{N}} \hat{a}_n$, suppose for the sake of contradiction that $\hat{a} \setminus \bigcup_{n \in \mathbb{N}} \hat{a}_n \neq \emptyset$. Since $\bigcup_{n \in \mathbb{N}} \hat{a}_n$ is closed, $\hat{a} \setminus \bigcup_{n \in \mathbb{N}} \hat{a}_n = \hat{a} \cap (\bigcup_{n \in \mathbb{N}} \hat{a}_n)^c$ is open, so there exists a nonzero base element $\hat{b} \subseteq \hat{a} \setminus \bigcup_{n \in \mathbb{N}} \hat{a}_n$. Then for all $n$, $\hat{a}_n \subseteq \hat{a} \setminus \hat{b} = \hat{a} \setminus b$, so by 4.11, $a_n \leq a \setminus b$, which makes $a \setminus b$ an upper bound for $\langle a_n \rangle$. Since $b < a = \sup_{n \in \mathbb{N}} a_n$, we have a contradiction. Thus $\bigcup_{n \in \mathbb{N}} \hat{a}_n = \hat{a}$. $\square$

Let $X$ be a topological space. Recall that $X$ is a Baire space if the intersection of every countable collection of dense open sets in $X$ is dense, and that $X$ is a Hausdorff space if distinct points in $X$ have disjoint open neighborhoods. In the following lemmas, we will use a weakened version of the Baire category theorem: If $X$ is a compact Hausdorff space, then $X$ is a Baire space.
4.15. **Lemma.** In a compact Hausdorff space, if a set is meager and open, then it is empty.

**Proof.** Let $X$ be a compact Hausdorff space. By the Baire category theorem, $X$ is a Baire space. Let $E \subseteq H$ be open and meager. Since $E$ is meager, we can write $E$ as the union of countably many nowhere dense sets $\{M_n\}_{n \in \mathbb{N}}$. Fix $n \in \mathbb{N}$. By definition, $\text{int}(\overline{M_n}) = \emptyset$, so $X = X \setminus \text{int}(\overline{M_n}) = X \setminus M_n = \text{int}(X \setminus M_n)$. Thus $\text{int}(X \setminus M_n)$ is dense in $X$. Also $\text{int}(X \setminus M_n)$ is open, so since $X$ is a Baire space, $\bigcap_{n \in \mathbb{N}} \text{int}(X \setminus M_n)$ is dense. Then $\bigcap_{n \in \mathbb{N}} \text{int}(X \setminus M_n) \subseteq X \setminus \bigcup_{n \in \mathbb{N}} M_n = X \setminus E$. Since $E$ is open, $X \setminus E$ is closed, so $X = \bigcap_{n \in \mathbb{N}} \text{int}(X \setminus M_n) \subseteq X \setminus E = X \setminus E$. Hence $E$ is empty. \qed

4.16. **Lemma.** In any topological space $Z$, the family $\mathcal{M}$ of meager subsets of $Z$ forms a $\sigma$-ideal.

**Proof.** Let $Z$ be a topological space and let $\mathcal{M}$ be the family of meager subsets of $X$. Since $\text{int}\emptyset = \emptyset$, $\emptyset \in \mathcal{M}$. Let $A \in \mathcal{M}$ and $B \subseteq A$. Let $\langle A_n \rangle_{n \in \mathbb{N}}$ be a sequence of subsets of $Z$ such that $\bigcup_{n \in \mathbb{N}} A_n = A$ and for all $n$, $\text{int}\overline{A_n} = \emptyset$. Then for all $n$, $\text{int}B \cap \overline{A_n} \subseteq \text{int}\overline{A_n} = \emptyset$. Since $\bigcup_{n \in \mathbb{N}} B \cap A_n = B \cap \bigcup_{n \in \mathbb{N}} A_n = B \cap A = B$, $B \in \mathcal{M}$. Now let $\langle A_n \rangle_{n \in \mathbb{N}} \in \mathcal{M}^\sigma$. For each $n$, there exists $\langle A^i_n \rangle_{i \in \mathbb{N}}$ a sequence of subsets of $Z$ such that $\bigcup_{i \in \mathbb{N}} A^i_n = A_n$ and for all $i$, $\text{int}\overline{A^i_n} = \emptyset$. Since $\langle A^i_n \rangle_{n,i \in \mathbb{N}}$ is countable, we can re-index it to a sequence $\langle B_k \rangle_{k \in \mathbb{N}}$. Then for all $k$, $B_k$ is nowhere dense, so $\bigcup_{n \in \mathbb{N}} A_n = \bigcup_{i,n \in \mathbb{N}} A^i_n = \bigcup_{k \in \mathbb{N}} B_k \in \mathcal{M}$. \qed

The following theorem is the Loomis-Sikorski representation of a $\sigma$-complete Boolean algebra. It will be used in the proof of the Stone representation theorem of a measure algebra in 4.20.

4.17. **Theorem.** (314M in [9]) Let $A$ be a $\sigma$-complete Boolean algebra, and $Z$ its Stone space. Let $\mathcal{E} = \{\hat{a} : a \in A\}$, and let $\mathcal{M}$ be the $\sigma$-ideal of meager subsets of $Z$. Then
\[ \Sigma = \{ E \triangle A : E \in \mathcal{E}, A \in \mathcal{M} \} \] is a \( \sigma \)-algebra of subsets of \( Z \), \( \mathcal{M} \) is a \( \sigma \)-ideal of \( \Sigma \), and \( A \) is isomorphic, as Boolean algebra, to \( \Sigma / \mathcal{M} \).

**Proof.** First, I will show that \( \Sigma \) is a \( \sigma \)-algebra and \( \mathcal{M} \) is a \( \sigma \)-ideal of \( \Sigma \). Since \( \emptyset = \hat{0} \in \mathcal{E} \), \( \emptyset \triangle \emptyset \in \Sigma \). Let \( F \in \Sigma \). There exist \( a \in B \), \( A \in \mathcal{M} \) such that \( F = \hat{a} \triangle A \). Since \( (\hat{a})^c = \hat{a}^c \in \mathcal{E} \),

\[
F^c = (\hat{a} \triangle A)^c \\
= ((\hat{a} \cap A^c) \cup (\hat{a}^c \cap A))^c \\
= (\hat{a} \cap A^c)^c \cap (\hat{a}^c \cap A)^c \\
= (\hat{a}^c \cup A) \cap (\hat{a} \cup A^c) \\
= ((\hat{a}^c \cup A) \cap A^c) \cup ((\hat{a}^c \cup A) \cap \hat{a}) \\
= [(\hat{a}^c \cap A^c) \cup (A \cap A^c)] \cup [(\hat{a}^c \cap \hat{a}) \cup (A \cap \hat{a})] \\
= (\hat{a}^c \cap A^c) \cup (A \cap \hat{a}) \\
= (\hat{a}^c \setminus A) \cup (A \setminus \hat{a}^c) \\
= \hat{a}^c \triangle A \in \Sigma.
\]

Now let \( \langle F_n \rangle_{n \in \mathbb{N}} \in \Sigma^\mathbb{N} \). For each \( n \), let \( a_n \in B \), \( A_n \in \mathcal{M} \) such that \( F_n = \hat{a}_n \triangle A_n \). Since \( B \) is \( \sigma \)-complete, \( a := \sup_{n \in \mathbb{N}} a_n \in B \). By 4.14, \( \bigcup_{n \in \mathbb{N}} \hat{a}_n = \hat{a} \).

Let \( C : = \hat{a} \setminus \bigcup_{n \in \mathbb{N}} \hat{a}_n \). I will show that \( C \) has empty interior and is nowhere dense. Since each \( \hat{a}_n \) is open, \( \bigcup_{n \in \mathbb{N}} \hat{a}_n \) is open, so \( C = \hat{a} \cap (\bigcup_{n \in \mathbb{N}} \hat{a}_n)^c \) is closed. Since \( \bigcup_{n \in \mathbb{N}} \hat{a}_n \) is open,

\[
\bigcup_{n \in \mathbb{N}} \hat{a}_n = \text{int}(\bigcup_{n \in \mathbb{N}} \hat{a}_n), \text{ so } \partial \bigcup_{n \in \mathbb{N}} \hat{a}_n = \text{cl}(\bigcup_{n \in \mathbb{N}} \hat{a}_n) \setminus \text{int}(\bigcup_{n \in \mathbb{N}} \hat{a}_n) = \text{cl}(\bigcup_{n \in \mathbb{N}} \hat{a}_n) \setminus \bigcup_{n \in \mathbb{N}} \hat{a}_n = C. \text{ Suppose } C \text{ contains an open set } U. \text{ Then } U \text{ and } \bigcup_{n \in \mathbb{N}} \hat{a}_n \text{ are disjoint. Since}
\]
Thus $x / \in F_n \setminus \hat{a}$. Then there exists $n \in \mathbb{N}$ such that $x \in F_n = \hat{a}_n \Delta A_n$. In other words, $x \in \hat{a}_n \cup A_n$ and $x \notin \hat{a}_n \cap A_n$. Since $\hat{a}_n \subseteq \hat{a}$ and $x \notin \hat{a}$, we must have $x \in A_n \subseteq C \cup (\bigcup_{n \in \mathbb{N}} A_n)$. Next, suppose $x \in \hat{a} \setminus \bigcup_{n \in \mathbb{N}} F_n$. I will show that if $x \notin \bigcup_{n \in \mathbb{N}} A_n$, then $x \in C$, so suppose $x \notin \bigcup_{n \in \mathbb{N}} A_n$. Since $x \notin \bigcup_{n \in \mathbb{N}} F_n = \bigcup_{n \in \mathbb{N}} \hat{a}_n \Delta A_n$, for all $n$ either $x \in \hat{a}_n \cap A_n$ or $x \notin \hat{a}_n \cup A_n$. By assumption, $x \notin \hat{a}_n \cap A_n$, so $x \notin \hat{a}_n \cup A_n$. Thus $x \notin \bigcup_{n \in \mathbb{N}} \hat{a}_n$, so $x \in \hat{a} \setminus \bigcup_{n \in \mathbb{N}} \hat{a}_n = C$. Thus $A \subseteq C \cup (\bigcup_{n \in \mathbb{N}} A_n)$. Since $\mathcal{M}$ is a $\sigma$-ideal, $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{M}$, so since $C$ is nowhere dense, $C \cup (\bigcup_{n \in \mathbb{N}} A_n) \in \mathcal{M}$. Finally, since $A \subseteq C \cup (\bigcup_{n \in \mathbb{N}} A_n)$, $A \in \mathcal{M}$. Thus, $\bigcup_{n \in \mathbb{N}} F_n = \emptyset \Delta (\bigcup_{n \in \mathbb{N}} F_n) = (\hat{a} \Delta \hat{a}) \Delta (\bigcup_{n \in \mathbb{N}} F_n) = \hat{a} \Delta A \in \Sigma$, so $\Sigma$ is closed under countable unions. Hence, $\Sigma$ is a $\sigma$-algebra. Since $\emptyset = \emptyset \in \mathcal{E}$, $\mathcal{M} = \{\emptyset \Delta A : A \in \mathcal{M}\} \subseteq \Sigma$, so by 4.16 $\mathcal{M}$ is a $\sigma$-ideal of $\Sigma$.

Next, let $\sim$ be a relation on $\Sigma$ such that for all $F_1, F_2 \in \Sigma$,

$$F_1 \sim F_2 \iff \exists K \in \mathcal{M} \text{ such that } F_1 = F_2 \Delta K.$$ 

I will show that $\sim$ is an equivalence relation. Let $F_1, F_2, F_3 \in \Sigma$. Since $F_1 = F_1 \Delta \emptyset$ and $\emptyset \in \mathcal{M}$, $\sim$ is reflexive. Suppose $F_1 = F_2 \Delta K$ for some $K \in \mathcal{M}$. Then $F_1 \Delta K = (F_2 \Delta K) \Delta K = F_2 \Delta \emptyset = F_2$, so $\sim$ is symmetric. Finally, let $J, K \in \mathcal{M}$ such that $F_1 = F_2 \Delta K$ and $F_2 = F_3 \Delta J$. Then $F_1 = F_2 \Delta K = (F_3 \Delta J) \Delta K = F_3 \Delta (J \Delta K)$ and since $J \Delta K \in \mathcal{M}$, $\sim$ is transitive.

Denote the equivalence class of $F \in \Sigma$ under $\sim$ by $[F]_{\mathcal{M}}$ and let $\Sigma/\mathcal{M} = \{[F]_{\mathcal{M}} : F \in \Sigma\}$. Define the operations $\triangle_{\mathcal{M}}$ and $\wedge_{\mathcal{M}}$ on $\Sigma/\mathcal{M}$ as follows: for all $f_1 = [F_1]_{\mathcal{M}}, f_2 = [F_2]_{\mathcal{M}}$.

$U$ is open and in the boundary of $\bigcup_{n \in \mathbb{N}} \hat{a}_n$. $U$ and $\bigcup_{n \in \mathbb{N}} \hat{a}_n$ has nonempty intersection, contradiction. Hence $\text{int}(C) = \emptyset$. Since $C$ is closed, $C = \overline{C}$, so $\text{int}(\overline{C}) = \emptyset$, i.e. $C$ is nowhere dense.
When context distinguishes between $\Delta$ and $\Delta_{\mathcal{M}}$, and $\land$ and $\land_{\mathcal{M}}$, we will drop the subscript $\mathcal{M}$. I will now show that $\Delta_{\mathcal{M}}$ and $\land_{\mathcal{M}}$ are well-defined, making $\sim$ a congruence relation. Let $F_1, F_2, G_1, G_2 \in \Sigma$ and suppose $F_1 \sim G_1$ and $F_2 \sim G_2$. Then there exist $K_1, K_2 \in \mathcal{M}$ such that $F_1 = G_1 \Delta K_1$ and $F_2 = G_2 \Delta K_2$, so $F_1 \triangle F_2 = (G_1 \Delta K_1) \Delta (G_2 \Delta K_2) = (G_1 \Delta G_1) \Delta (K_1 \Delta K_2)$. Since $K_1 \Delta K_2 \in \mathcal{M}$, $F_1 \triangle G_1 \sim G_1 \Delta G_2$. Also $F_1 \land F_2 = (G_1 \Delta K_1) \land (G_2 \Delta K_2) = (G_1 \land (G_2 \Delta K_2)) \Delta (K_1 \land (G_2 \Delta K_2)) = ((G_1 \land G_2) \Delta (G_1 \land K_2)) \Delta ((K_1 \land G_2) \Delta (K_1 \Delta K_2))$. Since $(G_1 \land K_2) \Delta (K_1 \land G_2) \Delta (K_1 \Delta K_2) \in \mathcal{M}$, $F_1 \land G_1 \sim G_1 \land G_2$.

Let $\phi : A \to \Sigma / \mathcal{M}$ such that $\phi(a) = [\hat{a}]_{\mathcal{M}}$. I will show that $\phi$ is a Boolean homomorphism. Let $a, b \in A$. Then

$$
\phi(a \triangle b) = [\hat{a} \triangle \hat{b}]_{\mathcal{M}}
$$

$$
= [\hat{a} \land \hat{b}]_{\mathcal{M}}
$$

$$
= \{(\hat{a} \land \hat{b}) \triangle K : K \in \mathcal{M}\}
$$

$$
= \{(\hat{a} \triangle K_1) \land (\hat{b} \triangle K_2) : K_1, K_2 \in \mathcal{M}\}
$$

$$
= \{x \triangle y : x \in [\hat{a}]_{\mathcal{M}}, y \in [\hat{b}]_{\mathcal{M}}\}
$$

$$
= [\hat{a}]_{\mathcal{M}} \land [\hat{b}]_{\mathcal{M}}
$$

$$
= \phi(a) \land \phi(b),
$$

$$
\phi(a \land b) = [\hat{a} \land \hat{b}]_{\mathcal{M}}
$$

$$
= [\hat{a} \land \hat{b}]_{\mathcal{M}}
$$

$$
= \{(\hat{a} \land \hat{b}) \triangle K : K \in \mathcal{M}\}
$$
\[
\begin{align*}
&= \{(\hat{a} \cap \hat{b}) \Delta (\hat{a} \cap K_2) \Delta (\hat{b} \cap K_1) \Delta (K_1 \cap K_2) : K_1, K_2 \in \mathcal{M}\} \\
&= \{(\hat{a} \Delta K_1) \cap (\hat{b} \Delta K_2) : K_1, K_2 \in \mathcal{M}\} \text{ by 4.1.4} \\
&= \{x \cap y : x \in [\hat{a}]_\mathcal{M}, y \in [\hat{b}]_\mathcal{M}\} \\
&= [\hat{a}]_\mathcal{M} \cap [\hat{b}]_\mathcal{M} \\
&= \phi(a) \cap \phi(b), \\
\phi(1) &= [\hat{1}]_\mathcal{M} \\
&= 1_{\Sigma/\mathcal{M}} 
\end{align*}
\]

Finally, I will show that \( \mathcal{A} \cong \Sigma/\mathcal{M} \) as Boolean algebras. Let \( F \in \Sigma \). By definition, \( F \) is expressible as \( E \Delta A \) for some \( E \in \mathcal{E} \) and \( A \in \mathcal{M} \), so \( \phi \) is surjective. Note that \( F \Delta E = (E \Delta A) \Delta E = (A \Delta E) \Delta E = \Lambda \Delta (E \Delta E) = \Lambda \Delta \emptyset = A \in \mathcal{M} \). To see that \( E \) is the unique element of \( \mathcal{E} \) such that \( F \Delta E \in \mathcal{M} \), let \( E' \in \mathcal{E} \) such that \( E \neq E' \). By 4.13, \( E \) and \( E' \) are open, so \( E \Delta E' \) is open. Let \( x \in E \Delta E' \). If \( x \in A \), then certainly \( x \in A \cup (F \Delta E') \). If \( x \notin A \), then \( x \in (E \Delta E') \setminus A \subseteq (E \Delta E') \Delta A = (E \Delta A) \Delta E' = F \Delta E' \subseteq A \cup (F \Delta E') \).

Thus \( E \Delta E' \subseteq A \cup (F \Delta E') \). By 4.15, \( E \Delta E' \) is not meager, so \( A \cup (F \Delta E') \notin \mathcal{M} \). Since \( A \in \mathcal{M} \), we have \( F \Delta E' \notin \mathcal{M} \). Hence \( E \) is unique, so \( \phi \) is injective. Thus \( \phi \) is a bijection. Since \( \phi \) is also a Boolean homomorphism, it is an isomorphism, so \( \mathcal{A} \cong \Sigma/\mathcal{M} \). \( \square \)

4.18. Definition. (321A in [9]) A measure algebra is a pair \((\mathcal{A}, \nu)\) where \( \mathcal{A} \) is a \( \sigma \)-complete Boolean algebra and \( \nu : \mathcal{A} \to [0, \infty] \) is a function such that

1. \( \nu(0) = 0 \)

2. whenever \( (a_n)_{n \in \mathbb{N}} \) is a disjoint sequence in \( \mathcal{A} \) \( (a_i \land a_j = 0 \) for distinct \( i, j \)),

\[ \nu(\sup_{n \in \mathbb{N}} a_n) = \sum_{n=0}^{\infty} \nu(a_n) \]
(3) $\nu$ is strictly positive ($a = 0 \iff \nu(a) = 0$)

Note that probability algebras are measure algebras with $\nu(1) = 1$.

Given measure algebras $(A, \nu)$ and $(B, \mu)$, a function $\phi : A \to B$ is a measure algebra isomorphism if $\phi$ is a Boolean algebra isomorphism that also satisfies $\nu(E) = \mu(\phi(E))$ for all $E \in A$.

**Definition.** (112A in [8]) A measure space is a triple $(X, \Sigma, \mu)$ where

1. $X$ is a set
2. $\Sigma$ is a $\sigma$-algebra of subsets of $X$
3. $\mu : \Sigma \to [0, \infty]$ is a function such that
   a. $\mu(\emptyset) = 0$
   b. if $\langle E_n \rangle_{n \in \mathbb{N}}$ is a disjoint sequence in $\Sigma$, then $\mu(\bigcup_{n \in \mathbb{N}} E_n) = \sum_{n=0}^{\infty} \mu E_n$.

The measure algebra of a measure space $(X, \Sigma, \mu)$ is the construction $(\hat{B}, \hat{\mu})$ in 4.3, which was shown to be a measure algebra.

**Definition.** (313H in [9]) Let $P, Q$ be posets. A function $\theta : P \to Q$ is sequentially order-continuous if the following are true:

1. for all $p_1, p_2 \in P$, $p_1 \leq p_2 \implies \theta(p_1) \leq \theta(p_2)$
2. if $\langle p_n \rangle_{n \in \mathbb{N}}$ is a non-decreasing sequence in $P$ and $p = \sup_n p_n$, then $\theta(p) = \sup_n \theta(p_n)$

4.19. **Proposition.** Let $\leq_\Sigma$ be a binary relation on $\Sigma$ such that for all $F_1, F_2 \in \Sigma$, $F_1 \leq_\Sigma F_2 \iff E_1 \subseteq E_2$, where $E_1, E_2 \in \mathcal{E}$ such that $F_i = E_i \triangle A_i$ for some $A_1, A_2 \in \mathcal{M}$. $\leq_\Sigma$ is a partial order.
Proof. Let $F_1, F_2 \in \Sigma$ and $E_1, E_2 \in \mathcal{E}$ such that $F_i = E_i \Delta A_i$ for some $A_1, A_2 \in \mathcal{M}$. There exist $a_1, a_2 \in \mathcal{A}$ such that $\hat{a}_i = E_i$. Since $E_1 \subseteq E_2 \iff a_1 \leq a_2$, $F_1 \leq_{\Sigma} F_2 \iff a_1 \leq a_2$. Since $\leq$ is a partial order on $\mathcal{A}$, it follows immediately that $\leq_{\Sigma}$ is a partial order on $\Sigma$. \qed

The next theorem describes the Stone representation of a measure algebra.

4.20. Theorem. (321J in [9]) Let $(\mathcal{A}, \hat{\mu})$ be a measure algebra. Then there exists a measure space whose measure algebra is isomorphic, as a measure algebra, to $(\mathcal{A}, \hat{\mu})$.

Proof. As a Boolean algebra, $\mathcal{A}$ is isomorphic to $\Sigma/\mathcal{M}$ as constructed in 314M. Denote the canonical isomorphism by $\pi : \Sigma/\mathcal{M} \to \mathcal{A}$, $\pi([\hat{a}]_{\mathcal{M}}) = a$. Let $\theta : \Sigma \to \mathcal{A}$ such that $\theta(F) = \pi([F]_{\mathcal{M}})$ for all $F \in \Sigma$. I will show that $\theta$ is a sequentially order-continuous surjective Boolean homomorphism with kernel $\mathcal{M}$. Let $A, B \in \Sigma$. Then $\theta(A \Delta B) = \pi([A \Delta B]_{\mathcal{M}}) = \pi([A]_{\mathcal{M}} \Delta [B]_{\mathcal{M}}) = \pi([A]_{\mathcal{M}}) \Delta \pi([B]_{\mathcal{M}}) = \theta(A) \Delta \theta(B)$. Similarly, $\theta(A \cap B) = \theta(A) \wedge \theta(B)$, so $\theta$ is a Boolean homomorphism. The kernel of $\theta$ is $\mathcal{M}$ because for all $F \in \Sigma$, $\theta(F) = 0 \iff \pi([F]_{\mathcal{M}}) = 0 \iff [F]_{\mathcal{M}} = [\emptyset]_{\mathcal{M}} \iff F \sim \emptyset \iff \exists K \in \mathcal{M} F = \emptyset \Delta K = K \iff F \in \mathcal{M}$. Since $\pi$ and $F \mapsto [F]_{\mathcal{M}}$ are surjective, $\theta$ is surjective. To show that $\theta$ is sequentially order-continuous, let $F_1, F_2 \in \Sigma$ such that $F_1 \leq_{\Sigma} F_2$. Since $\pi$ is an isomorphism, there exist $a_1, a_2 \in \mathcal{A}$ such that $[F_i]_{\mathcal{M}} = [\hat{a}_i]_{\mathcal{M}} = \{\hat{a}_i \Delta K : K \in \mathcal{M}\}$ for $i = 1, 2$. Then there exist $A_1, A_2 \in \mathcal{M}$ such that $F_i = \hat{a}_i \Delta A_i$ for $i = 1, 2$. Since $F_1 \leq_{\Sigma} F_2$, $\theta(F_1) = \pi([F_1]_{\mathcal{M}}) = \pi([\hat{a}_1]_{\mathcal{M}}) = a_1 \leq a_2 = \pi([\hat{a}_2]_{\mathcal{M}}) = \pi([F_2]_{\mathcal{M}}) = \theta(F_2)$.

Now let $(F_n)_{n \in \mathbb{N}}$ be a non-decreasing sequence in $\Sigma$. Suppose $F = \sup_n F_n \in \Sigma$ (if the sup does not exist, then there is nothing to show). For each $n$, let $E_n \in \mathcal{E}$, $A_n \in \mathcal{M}$, $a_n \in \mathcal{A}$ such that $F_n = E_n \Delta A_n$ and $\hat{a}_n = E_n$. Also let $E \in \mathcal{E}$, $A \in \mathcal{M}$, $a \in \mathcal{A}$ such that
\[ F = E \triangle A \text{ and } \hat{a} = E. \] Then

\[ F = \sup_n F_n \iff \forall n \ F_n \leq F \text{ and } \forall F' \in \Sigma \text{ if } \forall n \ F_n \leq F', \text{ then } F \leq F' \]

\[ \iff \forall n \ E_n \subseteq E \text{ and } \forall E' \in \mathcal{E} \text{ if } \forall n \ E_n \subseteq E', \text{ then } E \subseteq E' \]

\[ \iff \forall n \ a_n \leq a \text{ and } \forall a' \in A \text{ if } \forall n \ a_n \leq a', \text{ then } a \leq a' \]

\[ \iff a = \sup_n a_n, \]

so \( \theta(F) = \pi([F]_\mathcal{M}) = \pi([\hat{a}]_\mathcal{M}) = a = \sup_n a_n = \sup_n \pi([\hat{a}_n]_\mathcal{M}) = \sup_n \pi([F_n]_\mathcal{M}) = \sup_n \theta(F_n) \). Hence, \( \theta \) is sequentially order-continuous.

Let \( \nu : \Sigma \to [0, \infty) \) such that \( \nu(E) = \hat{\mu}(\theta(E)) \) for all \( E \in \Sigma \). I will show that this makes \((Z, \Sigma, \nu)\) a measure space. First, by 314M \( \Sigma \) is a \( \sigma \)-algebra. Next, \( \nu(\emptyset) = \hat{\mu}(\theta(\emptyset)) = \hat{\mu}(\pi([\emptyset]_\mathcal{M})) = \hat{\mu}(\pi([\hat{0}]_\mathcal{M})) = \hat{\mu}(0) = 0 \). Now let \( \langle F_n \rangle_{n \in \mathbb{N}} \) be a disjoint sequence in \( \Sigma \), i.e. for all \( i \neq j \), \( F_i \cap F_j = \emptyset \). Let \( i \neq j \in \mathbb{N} \). Since \( \theta \) is a Boolean homomorphism, \( \theta(F_i) \land \theta(F_j) = \theta(F_i \cap F_j) = \theta(\emptyset) = 0 \), so \( \langle \theta(F_n) \rangle_{n \in \mathbb{N}} \) is a disjoint sequence. For each \( n \), let \( E_n \in \mathcal{E}, A_n \in \mathcal{M}, a_n \in A \) such that \( F_n = E_n \triangle A_n \) and \( \hat{a}_n = E_n \). Let \( i \neq j \in \mathbb{N} \). First we show that \( E_i \cap E_j \in \mathcal{M} \). By 4.1.4,

\[
\emptyset = F_i \cap F_j = (E_i \triangle A_i) \cap (E_j \triangle A_j) \\
= (E_i \cap (E_j \triangle A_j)) \triangle (A_i \cap (E_j \triangle A_j)) \\
= [(E_i \cap E_j) \triangle (E_i \cap A_j)] \triangle (A_i \cap (E_j \triangle A_j)) \\
= (E_i \cap E_j) \triangle [(E_i \cap A_j) \triangle (A_i \cap (E_j \triangle A_j))].
\]
Since $\mathcal{M}$ is a $\sigma$-ideal, $E_i \cap A_j, A_i \cap (E_j \triangle A_j) \in \mathcal{M}$. Furthermore, $(E_i \cap A_j) \triangle (A_i \cap (E_j \triangle A_j)) \subseteq (E_i \cap A_j) \cup (A_i \cap (E_j \triangle A_j)) \in \mathcal{M}$, so $(E_i \cap A_j) \triangle (A_i \cap (E_j \triangle A_j)) \in \mathcal{M}$. So we have found $K \in \mathcal{M}$ such that $(E_i \cap E_j) \triangle K = \emptyset$, which implies $E_i \cap E_j = (E_i \cap E_j) \triangle \emptyset = (E_i \cap E_j) \triangle (K \triangle K) = ((E_i \cap E_j) \triangle K) \triangle K = \emptyset \triangle K = K \in \mathcal{M}$. Using this and the fact that $\pi$ is a Boolean homomorphism, we get

$$0 = \pi([\emptyset]_{\mathcal{M}})$$

$$= \pi([E_i \cap E_j]_{\mathcal{M}})$$

$$= \pi([E_i]_{\mathcal{M}} \cap_{\mathcal{M}} [E_j]_{\mathcal{M}})$$

$$= \pi([E_i]_{\mathcal{M}}) \land \pi([E_j]_{\mathcal{M}})$$

$$= a_i \land a_j,$$

so $\langle a_n \rangle_{n \in \mathbb{N}}$ is a disjoint sequence. Since $\theta$ is sequentially order-continuous, $\theta(\cup_{n \in \mathbb{N}} F_n) = \sup_n \theta(F_n)$. Hence

$$\nu(\cup_{n \in \mathbb{N}} F_n) = \hat{\mu}(\theta(\cup_{n \in \mathbb{N}} F_n)) = \hat{\mu}(\sup_n \theta(F_n)) = \sum_{n=0}^{\infty} \hat{\mu}(F_n) = \sum_{n=0}^{\infty} \nu(F_n)$$

so $(Z, \Sigma, \nu)$ a measure space.

For all $E \in \Sigma$,

$$\nu(E) = 0 \iff \hat{\mu}(\theta(E)) = 0 \iff \theta(E) = 0 \iff E \in \ker(\theta) = \mathcal{M}$$

so $\mathcal{M}$ consists of exactly the sets of measure zero in $(Z, \Sigma, \nu)$. Thus the measure algebra of $(Z, \Sigma, \nu)$ is $(\Sigma/\mathcal{M}, \hat{\nu})$, where $\hat{\nu}([E]_{\mathcal{M}}) = \nu(E)$. Since the Boolean algebra isomorphism $\pi : \Sigma/\mathcal{M} \to \mathcal{A}$ also satisfies $\hat{\nu}([E]_{\mathcal{M}}) = \nu(E) = \hat{\mu}(\theta(E)) = \hat{\mu}(\pi([E]_{\mathcal{M}}))$, $\pi$ is a measure algebra isomorphism between $(\Sigma/\mathcal{M}, \hat{\nu})$ and $(\mathcal{A}, \hat{\mu})$. 

45
Notation. For all $a, b \in \mathbb{R}$, let

$$a \div b = \begin{cases} a - b & \text{if } a \geq b \\ 0 & \text{otherwise} \end{cases}$$

Definition. Let $Pr$ denote the following set of axioms.

(i) Boolean algebra axioms

1. $\sup_x \hat{d}(x \lor x, x)$
2. $\sup_x \hat{d}(x \land x, x)$
3. $\sup_x \sup_y \hat{d}(x \lor y, y \lor x)$
4. $\sup_x \sup_y \hat{d}(x \land y, y \land x)$
5. $\sup_x \sup_y \hat{d}((x \land y) \lor y, y)$
6. $\sup_x \sup_y \hat{d}((x \lor y) \land y, y)$
7. $\sup_x \sup_y \sup_z \hat{d}((x \lor y) \lor z, x \lor (y \lor z))$
8. $\sup_x \sup_y \sup_z \hat{d}((x \land y) \land z, x \land (y \land z))$
9. $\sup_x \sup_y \sup_z \hat{d}(x \lor (y \land z), (x \lor y) \land (x \lor z))$
10. $\sup_x \sup_y \sup_z \hat{d}(x \land (y \lor z), (x \land y) \lor (x \land z))$
11. $\sup_x \hat{d}(x \lor x^c, 1)$
12. $\sup_x \hat{d}(x \land x^c, 0)$

(ii) measure axioms

13. $\hat{\mu}(0)$
14. $\neg \hat{\mu}(1)$
15. $\sup_x \sup_y (\hat{\mu}(x \land y) \div \hat{\mu}(x))$
16. \( \sup_x \sup_y (\hat{\mu}(x) - \hat{\mu}(x \lor y)) \)

17. \( \sup_x \sup_y |(\hat{\mu}(x) - \hat{\mu}(x \land y)) - (\hat{\mu}(x \lor y) - \hat{\mu}(y))| \)

(iii) connection between \( \hat{d} \) and \( \hat{\mu} \)

18. \( \sup_x \sup_y |\hat{d}(x, y) - \hat{\mu}(x \triangle y)| \)

**Remark.** From axioms 15–17, we can show \( \hat{\mu}(x \land y) + \hat{\mu}(x \lor y) = \hat{\mu}(x) + \hat{\mu}(y) \) for all \( x, y \in \hat{\mathcal{B}} \): Let \( x, y \in \hat{\mathcal{B}} \). By axiom 15, \( \hat{\mu}(x \land y) - \hat{\mu}(x) = 0 \), so \( \hat{\mu}(x \land y) \leq \hat{\mu}(x) \). By axiom 16, \( \hat{\mu}(y) \leq \hat{\mu}(x \lor y) \). Thus, \( \hat{\mu}(x) - \hat{\mu}(x \land y) = \hat{\mu}(x) - \hat{\mu}(x \land y) = \hat{\mu}(x \lor y) - \hat{\mu}(y) = \hat{\mu}(x \lor y) - \hat{\mu}(y) \) by axiom 17.

Axioms 1–12 are direct translations of the first 12 axioms of a Boolean algebra. Conspicuously missing from (i) is last Boolean algebra axiom, \( 0 \neq 1 \). In the language of continuous logic, \( 0 \neq 1 \) informally translates to \( \hat{d}(0, 1) \neq 0 \); for \( 0 \neq 1 \) to hold in a model \( \mathcal{M} \), the truth value \( \mathcal{M}(\hat{d}(0, 1)) \) must be \( > 0 \). But \( \mathcal{M} \not\models \hat{d}(0, 1) \) does not imply that \( \mathcal{M} \models \neg \hat{d}(0, 1) \), since the latter occurs only when \( \mathcal{M}(\hat{d}(0, 1)) = 1 \). So while we can always express \( x = y \) in \( \mathcal{M} \) by \( \mathcal{M} \models d(x, y) \), we usually cannot express \( x \neq y \) [6]. However, by virtue of the additional structure imposed by the axioms of \( Pr \), we can not only show that \( \mathcal{M} \not\models \hat{d}(0, 1) \), but also that \( \mathcal{M} \models \neg \hat{d}(0, 1) \).

4.21. **Proposition.** \( Pr \models \neg \hat{d}(0, 1) \).

**Proof.** Let \( \mathcal{M} \) be a model of \( Pr \). By axiom 18,

\[
|\mathcal{M}(\hat{d}(0, 1)) - \mathcal{M}(\hat{\mu}(0 \triangle 1))|_E = \mathcal{M}(|\hat{d}(0, 1) - \hat{\mu}(0 \triangle 1)|) = 0,
\]

where \(|\cdot|_E \) denotes one-dimensional Euclidean distance. Thus \( \mathcal{M}(\hat{d}(0, 1)) \)

\[
= \mathcal{M}(\hat{\mu}(0 \triangle 1))
\]
\[
\hat{\mu}^M((0 \triangle 1)^M) \\
= \hat{\mu}^M(((0 \land 1^c) \lor (0^c \land 1))^M) \\
= \hat{\mu}^M((0^M \land (1^M)^c) \lor (0^M)^c \land 1^M) \\
= \hat{\mu}^M(0 \land 1^c) + \hat{\mu}^M(0^c \land 1) - \hat{\mu}^M((0 \land 1^c) \land (0^c \land 1)) \text{ by axioms 15 - 17} \\
\leq \hat{\mu}(0) + \hat{\mu}^M(0^c \land 1) + \hat{\mu}(0) \text{ by axiom 15} \\
= 0 + \hat{\mu}^M(0^c \land 1) + 0 \text{ since } \hat{\mu}^M(0^M) = \hat{\mu}^M((0 \land 1^c) \land (0^c \land 1)) = 0 \text{ by axiom 13} \\
= \hat{\mu}^M(0^c \land 1) \\
= \hat{\mu}^M(1 \land 1) \text{ since } 0^c = (0 \land 0^c)^c = 0^c \lor 0 = 1 \\
= \hat{\mu}^M(1) \text{ by axiom 1} \\
= 1 \text{ since } 1 - \hat{\mu}^M(1) = 1 - \hat{\mu}(1) = \hat{\mu}(\neg(1)) = 0.
\]

Then \(\mathcal{M}(\neg\hat{d}(0, 1)) = 1 - \mathcal{M}(\hat{d}(0, 1)) = 1 - 1 = 0\), so \(\mathcal{M} \models \neg\hat{d}(0, 1)\).

Now we show that \(Pr\) axiomatizes the class of probability structures.

4.22. Theorem. (16.1 from [2]) Let \(\mathcal{M}\) be an \(L_{Pr}\)-structure with underlying metric space \((\mathcal{M}, \hat{d})\). Then \(\mathcal{M}\) is a model of \(Pr\) if and only if \(\mathcal{M}\) is the probability structure associated with a probability space \((X, \Sigma, \mu)\) as above.

Proof. (\(\Rightarrow\)) Let \((X, \Sigma, \mu)\) be a probability space and let \(\mathcal{M} = (\hat{\mathcal{B}}, 0, 1, \land, \lor, \hat{\mu})\) be its associated probability structure. Let \(x = [A]_\mu, y = [B]_\mu, z = [C]_\mu \in \hat{\mathcal{B}}\). The proofs that each axiom of \(Pr\) is true in \(\mathcal{M}\) are similar, so I will just show one from each of (i)-(iii).

From the set identity \((A \cup (B \cap C)) = ((A \cup B) \cap (A \cup C))\), we get \(0 = \mu(0) = \mu((A \cup (B \cap C)) \triangle ((A \cup B) \cap (A \cup C))) = \hat{\mu}((x \lor (y \land z)) \triangle ((x \lor y) \land (x \lor z))) = \hat{d}(x \lor (y \land z), (x \lor y) \land
(x ∨ z)). Since x, y, and z were arbitrary, \( \sup_x \sup_y \sup_z \hat{d}(x \lor (y \land z), (x \lor y) \land (x \lor z)) = 0 \), so \( \mathcal{M} \models \text{axiom 9} \).

Note that \( \mu A \geq \mu(A \land B) \) and \( \mu(B \lor A) \geq \mu(B) \), so \( \mu(A) \land \mu(A \land B) = \mu(A) - \mu(A \land B) \) and \( \mu(A \land B) \lor \mu(B) = \mu(A \lor B) - \mu(B) \). Rearranging the set identity \( \mu(A) + \mu(B) = \mu(A \land B) + \mu(A \lor B) \) to \( \mu(A) - \mu(A \land B) = \mu(A \lor B) - \mu(B) \), we get \( 0 = (\mu(A) - \mu(A \lor B)) - (\mu(A \lor B) - \mu(B)) = (\mu(A) - \mu(A \land B)) - (\mu(A \lor B) - \mu(B)) \) since x and y were arbitrary, \( \sup_x \sup_y [(\mu(x) - \mu(x \land y)) - (\mu(x \lor y) - \mu(y))] = 0 \), so \( \mathcal{M} \models \text{axiom 17} \).

Finally, since \( \hat{d}(x, y) = \mu(x \triangle y), \sup_x \sup_y \hat{d}(x, y) - \mu(x \triangle y) = 0 \), so \( \mathcal{M} \models \text{axiom 18} \).

(\( \Rightarrow \)) Let \( \mathcal{M} \) be a model of \( \text{Pr} \). By assumption, \( \mathcal{M} \) is an \( L^{\text{Pr}} \)-structure, so it is of the form \((A, 0, 1, \land, \lor, \nu)\) for a set \( A \), distinguished constants 0 and 1, a unary function symbol \( \land \), two binary function symbols \( \land \) and \( \lor \), and a unary predicate symbol \( \nu \). Let \((A, d_\nu)\) be the underlying metric space of \( A \). Note that by definition, \((A, d_\nu)\) is complete. Since the axioms in (i) and \( 0 \neq 1 \) characterize a Boolean algebra, \( A \) is a Boolean algebra.

The measure axioms make \( \nu \) finitely additive. Let \( x, y \in A \). By the preceding remark, \( \nu(x) - \nu(x \land y) = \nu(x \lor y) - \nu(y) \). If \( x \land y = 0 \), then \( \nu(x) = \nu(x) - \nu(0) = \nu(x \lor y) - \nu(y) \), so \( \nu(x) + \nu(y) = \nu(x \lor y) \).

By (iii), the metric \( d_\nu \) must satisfy \( d_\nu(a, b) = \nu(a \triangle b) \). Given \( x \in A \), \( 1 \land x = (x^c \lor x) \land x = x \) by axioms 11 and 6. In other words, for all \( x \in A \), \( x \leq 1 \), so \( \nu(x) \leq \nu(1) = 1 \). Hence \( \nu \) is a finitely additive probability measure.

Now I will show that \( \nu \) is 1-Lipschitz on \( A \) with respect to \( d_\nu \). Let \( x, y \in A \). Since \( x = (x \land y^c) \lor (x \land y) = (x \land y) \lor (x \land y) \) and \( \nu \) is finitely additive, \( \nu(x) = \nu(x \land y) + \nu(x \land y) \). Similarly, \( \nu(y) = \nu(y \land x) + \nu(y \land x) \). Then \( |\nu(x) - \nu(y)| = |\nu(x \land y) +
\[ \nu(x \wedge y) - (\nu(y \setminus x) + \nu(y \wedge x)) = |\nu(x \setminus y) - \nu(y \setminus x)|. \] Since \( \nu(x \setminus y), \nu(y \setminus x) \geq 0 \),
\[ 0 \leq 2\nu(x \setminus y) \leq 2\nu(x \setminus y) + 2\nu(y \setminus x), \] so subtracting \( \nu(x \setminus y) + \nu(y \setminus x) \) from all sides,
\[ -(\nu(x \setminus y) + \nu(y \setminus x)) \leq \nu(x \setminus y) - \nu(y \setminus x) \leq \nu(x \setminus y) + \nu(y \setminus x). \] Thus
\[ |\nu(x) - \nu(y)| = |\nu(x \setminus y) - \nu(y \setminus x)| \leq \nu(x \setminus y) + \nu(y \setminus x) = \nu(x \triangle y) = d_\nu(x, y), \] so \( \nu \) is 1-Lipschitz.

Next, I will show that increasing sequence in \( \mathcal{A} \) is necessarily a Cauchy sequence with respect to \( d_\nu \). Let \( \langle x_n \rangle \in \mathcal{A}^\mathbb{N} \) be an increasing sequence. Recall that \( x_n \leq x_{n+1} \) means \( x_n \wedge x_{n+1} = x_n \). For the sake of contradiction, suppose \( \langle x_n \rangle \) is not Cauchy. Then there exists \( \epsilon > 0 \) such that for all \( N \in \mathbb{N} \) there exist \( m, n > N \) such that \( d_\nu(x_m, x_n) > \epsilon \).

Choose \( n_1 < n_2 \) such that \( d_\nu(x_{n_1}, x_{n_2}) > \epsilon \). Setting \( N_1 = n_2 + 1 \), choose \( N_1 < n_3 < n_4 \) such that \( d_\nu(x_{n_3}, x_{n_4}) > \epsilon \). In general, let \( N_k = n_{2k} + 1 \) and choose \( N_k < n_{2k+2} < n_{2k+3} \) such that \( d_\nu(x_{n_{2k+2}}, x_{n_{2k+3}}) > \epsilon \). For all \( k \in \mathbb{N} \), let \( y_k = x_{n_k} \). Note that for all \( i \in \mathbb{N} \),
\[ d_\nu(y_{2i+1}, y_{2i+2}) > \epsilon. \] Let \( 1/\epsilon < M \in \mathbb{N} \). Since \( \langle y_k \rangle \) is an increasing sequence, by 4.1.2
\[ \nu(y_1 \triangle y_{2M+1}) = d_\nu(y_1, y_{2M+1}) = \sum_{j=1}^{2M} d_\nu(y_j, y_{j+1}) \geq \sum_{j=0}^{M-1} d_\nu(y_{2j+1}, y_{2j+2}) > M\epsilon > 1, \] which is impossible, since 1 is the maximal element of \( \mathcal{A} \) and \( \nu(1) = 1 \). Therefore, \( \langle x_n \rangle \) is Cauchy with respect to \( d_\nu \). Since \( \langle \mathcal{A}, d_\nu \rangle \) is a complete metric space, for all increasing sequences \( \langle x_n \rangle \), there exists \( x \in \mathcal{A} \) such that \( x_n \to x \).

Now I will show that \( \mathcal{A} \) is \( \sigma \)-complete as a Boolean algebra. Let \( \langle a_n \rangle \in \mathcal{A}^\mathbb{N} \) be a sequence. For each \( n \geq 0 \), let \( b_n = \bigvee_{k=0}^{n} a_k \). By axiom 6, for all \( n, a_n \leq b_n \) and \( b_n \leq b_{n+1} \).

Since \( \langle b_n \rangle \) is increasing, there exists \( b \in \mathcal{A} \) such that \( b_n \to b \). First we show that \( b \) is an upper bound of \( \langle a_n \rangle \). For the sake of contradiction, suppose that there exists \( n \in \mathbb{N} \) such that \( b < b_n \), i.e. \( b \wedge b_n = b \) and \( b \neq b_n \). Since \( b \neq b_n \) and \( d_\nu \) is a metric, \( d_\nu(b, b_n) > 0 \).

Let \( \epsilon = d_\nu(b, b_n) \) and \( m \in \mathbb{N} \) such that \( m \geq n \). Since \( \langle b_n \rangle \) is increasing, \( b < b_n \leq b_m \).
By 4.1.2, \(d_\nu(b, b_m) = \nu(b \Delta b_m) = \nu(b_m \setminus b) = \nu(b_m \setminus b_n) + \nu(b_n \setminus b) \geq \nu(b_n \setminus b) = \epsilon,\)
so \(d_\nu(b, b_m) \geq \epsilon\) for all \(m \geq n,\) contradicting the fact that \(b = \lim_n b_n.\) Hence, for all \(n,\)
\(a_n \leq b_n \leq b,\) so \(b\) is an upper bound for \(\langle a_n \rangle.\)

To see that \(b\) is the least upper bound, let \(c \in \mathcal{A}\) be an upper bound for \(\langle a_n \rangle.\) Clearly
\(b_0 = a_0 \leq c.\) For any \(n \in \mathbb{N}\) such that \(b_n \leq c,\)
\(b_{n+1} = \bigvee_{k=0}^{n+1} a_k = b_n \lor a_n.\) Since
\(a_n \leq c,\) by 4.1.5 \(b_{n+1} = b_n \lor a_n \leq c,\) so \(c\) is also an upper bound for \(\langle b_n \rangle.\) Let \(\epsilon > 0.\)
Since \(b_n \to b,\) there exists \(N \in \mathbb{N}\) such that for all \(n > N,\)
\(d_\nu(b, b_n) < \epsilon.\) By 4.1.4,
\((b_n \land c) \Delta (b \land c) = (b \Delta b_n) \land c \leq b \Delta b_n.\) Then for all \(n > N,\)
\(d_\nu((b_n \land c), (b \land c)) = \nu((b_n \Delta b) \land c) \leq \nu(b \Delta b_n) = d_\nu(b, b_n) < \epsilon,\) so \(\langle b_n \land c \rangle\) converges
to \(b \land c.\) But \(c\) is an upper bound for \(\langle b_n \rangle,\) so \(\langle b_n \land c \rangle = \langle b_n \rangle.\) Hence
\(b = \lim_n b_n = \lim_n (b_n \land c) = b \land c,\)
i.e. \(b \leq c.\) Thus \(\sup_n a_n = b \in \mathcal{A}.\)

Now I will show that \(\nu\) is \(\sigma\)-additive on \(\mathcal{A}.\) Let \(\langle a_n \rangle_{n \in \mathbb{N}}\) be a disjoint sequence in \(\mathcal{A}.\)
Since \(\nu\) is finitely additive, \(\nu(\bigvee_{k=0}^n a_k) = \sum_{k=0}^n \nu(a_k)\) for all \(n,\) so \(\lim_{n \to \infty} \nu(\bigvee_{k=0}^n a_k) = \lim_{n \to \infty} \sum_{k=0}^n \nu(a_k) = \sum_{k=0}^\infty \nu(a_k).\) For each \(n,\) let \(b_n = \bigvee_{k=0}^n a_k.\) Since \(\mathcal{A}\) is \(\sigma\)-complete,
there exists \(c_0 \in \mathcal{A}\) such that \(c_0 = \bigvee_{k \in \mathbb{N}} a_k.\) Since \(\langle b_n \rangle\) is increasing, there exists \(c_1 \in \mathcal{A}\)
such that \(b_n \to c_1.\) By the proof of the \(\sigma\)-completeness of \(\mathcal{A}\) in 4.22, \(c_0 = c_1.\) Let \(c\) be their
common value. Since \(\nu\) is 1-Lipschitz, \(|\nu(b_n) - \nu(c)| \leq d_\nu(b_n, c)\) for all \(n.\) But \(b_n \to c,\) so
\(d_\nu(b_n, c) \to 0.\) Hence \(\lim_{n \to \infty} |\nu(b_n) - \nu(c)| = 0,\) so
\(\nu(\bigvee_{k \in \mathbb{N}} a_k) = \nu(c) = \lim_{n \to \infty} \nu(b_n) = \lim_{n \to \infty} \nu(\bigvee_{k=0}^n a_k) = \sum_{k=0}^\infty \nu(a_k).\)

Since \(\mathcal{A}\) is \(\sigma\)-complete and \(\nu\) is \(\sigma\)-additive, \((\mathcal{A}, \nu)\) is a measure algebra as defined in 4.3.

By theorem 4.20, it follows that \(\mathcal{M}\) is a probability algebra. \(\square\)
We end this section with a few facts about atomless probability spaces.

**Definition.** A probability space \((X, \mathcal{B}, \mu)\) is *atomless* if for all \(B \in \mathcal{B}\), if \(\mu(B) > 0\) then there exists \(B' \in \mathcal{B}\) such that \(B' \subseteq B\) and \(0 < \mu(B') < \mu(B)\).

Let \((X, \mathcal{B}, \mu)\) be an atomless probability space. Then for any \(B \in X\), \(\mu(B) = 0\), so \(X\) must be uncountable. We will use without proof the fact that \((X, \mathcal{B}, \mu)\) is atomless if and only if for all \(B \in \mathcal{B}\), \(\mu(B) > 0\) implies there exist \(B_1, B_2 \in \mathcal{B}\) such that \(B_1, B_2 \subseteq B\), \(B_1\) and \(B_2\) are disjoint, and \(\mu(B_1) = \mu(B_2) = \mu(B)/2\). The following axiom encodes this fact:

\[
\sup_x \inf_y |\hat{\mu}(x \land y) - \hat{\mu}(x \land y^c)|
\]

Let \(PrA\) be \(Pr\) together with (1). Then \(PrA\) axiomatizes the class of probability algebras associated with atomless probability spaces.
5. CONTINUOUS LOGIC VERSUS FIRST-ORDER LOGIC

This section consists of continuous analogues of some familiar meta-theorems of first-order logic. All of the definitions and results in this section are from [3].

**Definition.** A set of formulas \( \Sigma \) is *inconsistent* if for every formula \( \phi \), \( \Sigma \vdash \phi \). Otherwise \( \Sigma \) is *consistent*.

5.1. (4.5 in [3]) A set of formulas \( \Sigma \) is inconsistent if and only if there exists \( n \in \mathbb{N} \) such that \( \Sigma \vdash (\frac{1}{2})^n \).

**Definition.** A theory \( T \) is *complete* if there exists an \( L \)-structure \( M \) and assignment \( \sigma \) such that \( T = \text{Th}(M) \). Otherwise \( T \) is *incomplete*.

**Remark.** In first-order logic, an equivalent formulation of a theory \( T \) being complete is that for all sentences \( \phi \), \( T \vdash \phi \) or \( \vdash \neg \phi \). However, the addition of truth values beyond 0 and 1 means that there are now sentences \( \phi \) such that neither \( \phi \) nor \( \neg \phi \) have models, i.e. for all models \( M \) and all assignments \( \sigma \), \( M(\phi, \sigma), M(\neg \phi, \sigma) \neq 0 \). For example, the truth value of \( \frac{1}{2} \) and \( \frac{1}{2} \) in any model under any assignment is \( \frac{1}{2} \). If we required of a theory \( T \) that for all \( \phi \), \( T \vdash \phi \) or \( \vdash \neg \phi \), then by 5.1, \( T \) would be inconsistent. Such is the rationale for defining the property of completeness semantically.
One example of a complete theory is $Pr A$ (see 16.2 in [2]). Thus all probability structures associated with atomless probability spaces, such as examples 2 and 3 in 4.8, are elementarily equivalent. In contrast, $Pr$ is an incomplete theory, since otherwise, examples 1 and 2 from 4.8 would model the same sentences. But example 1 is not atomless, so only example 2 models $\sup_x \inf_y |\hat{\mu}(x \land y) - \hat{\mu}(x \land y^c)|$. Hence $Pr$ is incomplete. Ongoing work by Berenstein and Henson includes a characterization of complete extensions of $Pr$. All complete extensions other than $Pr A$ are obtained from the addition of axioms describing the measure of each atom.

5.2. Theorem. (Soundness) Let $L$ be a continuous signature and let $\Gamma$ be a set of $L$-formulas. Then for every $L$-formula $\phi$,

(1) $\Gamma \vdash \phi$ implies $\Gamma \models \phi$

(2) If $\Gamma$ is satisfiable, then $\Gamma$ is consistent.

5.3. Theorem. (Completeness) Let $L$ be a continuous signature and let $\Gamma$ be a set of $L$-formulas. If $\Gamma$ is consistent, then $\Gamma$ is satisfiable.

A corollary to the completeness theorem in continuous logic is the compactness theorem, obtained in the same way as the compactness theorem is obtained from completeness in first-order logic.

Corollary. (Compactness) Let $\Gamma$ be a set of $L$-formulas. If every finite subset $\Gamma_0$ of $\Gamma$ is satisfiable, then $\Gamma$ is satisfiable.

Proof. For the sake of contradiction, suppose every finite $\Gamma_0 \subseteq \Gamma$ is satisfiable but $\Gamma$ is not satisfiable. By 5.3, $\Gamma$ is inconsistent, so by 5.1, there exists $n \in \mathbb{N}$ such that $\Gamma \vdash (\frac{1}{2})^n$. Since
proofs are finite in length, there must be some finite $\Gamma_0 \subseteq \Gamma$ such that $\Gamma_0 \vdash (\frac{1}{2})^n$. Since $(\frac{1}{2})^n$ always has truth value $(\frac{1}{2})^n$, by 5.1, $\Gamma_0$ is inconsistent. By 5.2, $\Gamma_0$ is not satisfiable, contrary to assumption. \qed

**Remark.** Consider the following statements:

1. for every set of formulas $\Gamma$, if $\Gamma$ is consistent then $\Gamma$ is satisfiable
2. for every set of formulas $\Gamma$ and every formula $\phi$, $\Gamma \models \phi$ implies $\Gamma \vdash \phi$

In first-order logic, (1) and (2) are equivalent, but in continuous logic, (1) does not imply (2). Crucial in the proof that (1) implies (2) in first-order logic is the fact that, in first-order logic, if $\Gamma \nvdash \phi$, then $\Gamma \cup \{\neg \phi\}$ is consistent. But in continuous logic, if we take $\Gamma = \emptyset$ and $\phi = \frac{1}{2}$, then $\nvdash \frac{1}{2}$ but $\{\neg \frac{1}{2}\}$ is inconsistent by 5.1.

In fact, (2) fails in continuous logic. Consider the following counterexample. Let $\phi$ be the formula $\sup_x \sup_y d(x,y)$, which is true in an $L$-structure $\mathcal{M}$ if and only if $\mathcal{M}$ consists of one element. In general, the truth value of $\phi$ in $\mathcal{M}$ is the diameter of $\mathcal{M}$. Let

$$\Gamma = \{\phi \div (\frac{1}{2})^n : n \in \mathbb{N}\}$$

and suppose $(\mathcal{M}, \sigma) \models \Gamma$. Then for every $n \in \mathbb{N}$, $\max(\mathcal{M}(\phi, \sigma) - (\frac{1}{2})^n, 0) = \mathcal{M}(\phi \div (\frac{1}{2})^n, \sigma) = 0$, which means that for every $n \in \mathbb{N}$, $\mathcal{M}(\phi, \sigma) \leq (\frac{1}{2})^n$, so $\mathcal{M}(\phi, \sigma) = 0$. Thus $\Gamma \models \phi$. However, if $\Gamma \vdash \phi$, then there exists a finite $\Gamma_0 \subseteq \Gamma$ such that $\Gamma_0 \vdash \phi$, as proofs are finite in length. By 5.2, $\Gamma_0 \models \phi$. Let $N = \max(n : \phi \div (\frac{1}{2})^n \in \Gamma_0) < \infty$. Then any metric structure of diameter $\leq (\frac{1}{2})^N$ is a model of $\Gamma_0$ but not of $\phi$, so $\Gamma_0 \nvdash \phi$. Hence $\Gamma \nvdash \phi$.

There is, however, a counterpart to (2) in continuous logic; since (2) is sometimes called strong completeness, we will call its continuous version *approximated strong completeness*. 

55
5.4. **Corollary.** *(Approximated Strong Completeness)* Let \( L \) be a continuous signature, let \( \Gamma \) be a set of \( L \)-formulas, and let \( \phi \) be an \( L \)-formula. Then

\[
\Gamma \models \phi \iff \Gamma \vdash \phi \land (\frac{1}{2})^n
\]

for all \( n \in \mathbb{N} \).

**Definition.** Let \( \phi \) and \( \psi \) be \( L \)-formulas and let \( n \in \mathbb{N} \). Define \( n\psi \to \phi \) recursively:

1. \( 0\psi \to \phi := \phi \)
2. \( (n + 1)\psi \to \phi := \psi \to (n\psi \to \phi) \)

5.5. Among the meta-theorems of propositional logic and first-order logic is the deduction theorem: If \( \Gamma \) is a set of formulas and \( \phi \) and \( \psi \) are formulas, then

\[
\Gamma \cup \{\psi\} \vdash \phi \iff \Gamma \vdash \psi \to \phi.
\]

The left to right direction fails in continuous logic. Consider the following counterexample. Let \( \psi \) be \( \frac{1}{2} \) and let \( \phi \) be \( \frac{3}{4} \). Let \( \Gamma = \{\psi \to (\psi \to \phi)\} \). Then we have \( \psi \to (\psi \to \phi), \psi \vdash \phi \) via two applications of modus ponens. I will show, however, that \( \psi \to (\psi \to \phi) \not\models \psi \to \phi \). By the soundness theorem, it suffices to show \( \psi \to (\psi \to \phi) \not\models \psi \to \phi \). For any model \( M \) and assignment \( \sigma \), \( M(\psi, \sigma) = \frac{1}{2} \) and \( M(\phi, \sigma) = \frac{3}{4} \), so \( M(\psi \to \phi, \sigma) = \max(M(\phi, \sigma) - M(\psi, \sigma), 0) = \frac{1}{4} \). This means that \( \frac{1}{2} \to \frac{3}{4} \) has no models. On the other hand, \( M(\psi \to (\psi \to \phi), \sigma) = \max(M(\psi \to \phi, \sigma) - M(\psi, 0)) = \max(\frac{3}{4} - \frac{1}{2}, 0) = 0 \), so \( \frac{1}{2} \to (\frac{1}{2} \to \frac{3}{4}) \) is valid. Thus \( \psi \to (\psi \to \phi) \not\models \psi \to \phi \).

In particular, failure of the deduction theorem in continuous logic implies that the axiom

\[
(\alpha \to (\beta \to \gamma)) \to ((\alpha \to \beta) \to (\alpha \to \gamma))
\]
of classical propositional logic cannot be derived from the propositional part of continuous logic (axioms 1–6), since, together with axiom 1 and modus ponens, it implies the deduction theorem.

However, we do have the following weakened version of the deduction theorem in continuous logic.

5.6. **Theorem.** Let $\Gamma$ be a set of $L$-formulas and let $\phi, \psi$ be $L$-formulas. Then $\Gamma \cup \{\psi\} \vdash \phi$ if and only if for some $n \in \mathbb{N}$, $\Gamma \vdash n\psi \rightarrow \phi$.

Continuous logic satisfies a modified version of the Downward Löwenheim-Skolem theorem. Its statement requires the following definition.

**Definition.** Let $X$ be a topological space. The *density character of $X$*, denoted by $\|X\|$, is the minimum cardinality a dense subset of $X$.

5.7. **Theorem.** *(Downward Löwenheim-Skolem)* Let $\kappa$ be an infinite cardinal. Let $L$ be a signature such that $|L| \leq \kappa$. Let $M$ be an $L$-structure and let $A \subseteq M$ such that $\|A\| \leq \kappa$. Then there exists an elementary substructure $N$ of $M$ such that $A \subseteq N$ and $\|N\| \leq \kappa$.

This implies the Downward Löwenheim-Skolem of first-order logic (for any first-order language $L$, $L$-structure $M$, and cardinal $\kappa$ such that $\max(|L|, \aleph_0) \leq \kappa \leq |M|$, there exists an elementary substructure $N$ of $M$ such that $|N| = \kappa$) with cardinality of a model replaced by its density character: for any signature $L$, $L$-structure $M$, and cardinal $\kappa$ such that $\max(|L|, \aleph_0) \leq \kappa \leq \|M\|$, there exists an elementary substructure $N$ of $M$ such that $\|N\| = \kappa$. To see why, let $L$ be a signature, $M$ an $L$-structure, and $\kappa$ a cardinal such that
max(\(|L|, \aleph_0\)) \leq \kappa \leq \|M\|. Let \(A \subseteq M\) such that \(\|A\| = \kappa\). By 5.7, there exists an elementary substructure \(N\) of \(M\) such that \(A \subseteq N\) and \(\|N\| \leq \kappa\). Since \(\kappa = \|A\| \leq \|N\| \leq \kappa\), we have \(\|N\| = \kappa\). So in the same way that first-order logic cannot distinguish between infinite cardinalities, continuous logic cannot distinguish between density characters; if a theory \(T\) in continuous logic has a model, then \(T\) has a model of density character \(\leq \max(\|L\|, \aleph_0)\).

The first-order version of the Downward Löwenheim-Skolem theorem fails in continuous logic. This is witnessed by \(PrA\), whose signature is finite, but whose models are all uncountable, since all atomless probability spaces are uncountable.
6. Conclusion

Continuous logic was developed to study metric structures in a setting more convenient than that of first-order logic. Its model theory generalizes first-order model theory in the sense that any first-order structure is a metric structure if we equip it with the discrete metric. Restrictions of the connectives and quantifiers of continuous logic from $[0, 1]$ to \{0, 1\} yield those of first-order logic. The set of axioms from which a completeness result was proven in [3] are comparable to the axioms of first-order logic with equality, with the metric $d$ replacing $=$. Indeed, the equality axioms of first-order logic motivate the restriction of predicates and functions to be uniformly continuous.

Aside from its striking parallels to first-order logic in syntax, continuous logic satisfies modified versions of meta-theorems of first-order logic such as compactness, completeness, and Downward Löwenheim-Skolem. Yet while continuous logic is a useful tool to study the model theory of metric structures, it may not be indispensible. Clearly continuous logic is as strong as first-order logic, since every first-order formula has a direct translation in continuous logic, but is it stronger than first-order logic? Could the model theory of metric structures be done using first-order logic? Perhaps the two logics differ only in presentation and not in power of expression.

Lindström’s theorem cannot be used directly to compare these logics, since continuous logic is outside the scope of logics considered by Lindström’s theorem. To answer these questions, which are beyond the scope of this paper, one would first need to define the
notion of logic to include both first-order logic and continuous logic, then find a way to compare their expressive power.
REFERENCES


