On some aspects of measure and probability logics and a new logical proof for a theorem of Stone

Alireza Mofidi\textsuperscript{*a,b}

\textsuperscript{a}Department of Mathematics and Computer Science, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran
\textsuperscript{b}School of Mathematics, Institute for Research in Fundamental Sciences (IPM), P.O. Box: 19395-5746, Tehran, Iran

ABSTRACT: One of the functions of mathematical logic is studying mathematical objects and notions by logical means. There are several important representation theorems in analysis. Amongst them, there is a well-known classical one which concerns probability algebras. There are quite a few proofs of this result in the literature. This paper pursue two main goals. One is to consider some aspects of measure and probability logics and expose a novel proof for the mentioned representation theorem using ideas from logic and by application of an important result from model theory. The second and even more important goal is to present more connections between two fields of analysis and logic and reveal more the strength of logical methods and tools in analysis. The paper is mostly written for general mathematicians, in particular the people who are active in analysis or logic as the main audience. It is self-contained and includes all prerequisites from logic and analysis.

1. Introduction

There are several instances of applications of ideas and techniques from mathematical logic in analysis, probability theory, dynamical systems, etc. One can see for example [1, 2, 3, 5, 4, 6] as some instances of such interactions of logic. There are several important representation theorems in analysis and one of them concerns probability algebras. This result (see Theorem 3.6 below), due to Stone, is an important classical measure existence result. There are quite a few proofs for this theorem in the literature. During the course of our investigation in this paper, we mainly pursue two goals. One is to expose a relatively simple new proof for this representation theorem. The second goal is to elaborate the application of logical methods in analysis. In fact, our proof in a way indicates the power of logical methods in analysis and measure theory. The paper is self-contained and all prerequisites from logic and measure theory are included in it.

There are several ways for connecting measure and probability to logic. Our proof can be considered as an applications of the setting of the integration logic which is a framework for studying measure and probability structures through logical means. This setting was introduces in [3] and then was investigated in several works such as [1] and [4]. It is worth mentioning that in [4], this framework is represented as a specific example of a more abstract and general framework close to the functional analysis.

\textsuperscript{*}Corresponding author.
E-mail addresses: mofidi@aut.ac.ir
The technique of our proof uses the compactness theorem in logic which roughly states that if one has a family of mathematical properties (formally stated in a setting) and every finite number of them is satisfied in some structure, then there exists a structure which satisfies all of them together. Indeed, the method and strategy of the proof here by using this theorem seems to be more general than just the result of this paper and is possibly applicable, to some extent, in various measure existence results.

Presentation of the rest of the paper is as follows. In Section 2, we explain the logical setting and the technical statements we use in our proof. Section 3 contains the proof for the representation theorem we talked about on the base of the logical techniques.

2. Integration logic and some technical tools for constructing measures

First we recall that by a (Boolean algebra) measure on a Boolean algebra $\mathcal{B}$ of subsets of a set $M$, we mean a finitely additive real-valued function $\mu : \mathcal{B} \rightarrow [0, \infty]$ with the properties $\mu(\emptyset) = 0$ and that for any countable sequence $A_k \in \mathcal{B}$ of disjoint sets such that $\bigcup_k A_k \in \mathcal{B}$, we have $\mu(\bigcup_k A_k) = \sum_k \mu(A_k)$. If $\mathcal{B}$ is a $\sigma$-algebra, then we call $\mu$ a measure. As usual, if $f$ and $g$ are two real-valued functions, then we denote $\max(f, g)$ and $\min(f, g)$ by $f \vee g$ and $f \wedge g$ respectively.

Now we start to briefly review the logical settings of “integration logic”, investigated in [1], [2], [3] and [4] for studying measure and probability structures by logical methods. By a simple (measure relational) structure, (or simply, just a structure) in this paper, intuitively we mean a measure space equipped with a family of relations, where by a relation we mean a real-valued measurable function on the measure space or some power of it as its domain. We usually assign a symbol for each of such relations and call them relation symbols.

**Definition 2.1.** We also call the set of symbols mentioned above a (relational) language and usually denote it by $\mathcal{L}$.

We call the structure to which the symbols of $\mathcal{L}$ is referring a $\mathcal{L}$-structure (it will be formally defined in Definition 2.2). In order to do a systematic study on one or a family of structures by logical means, we usually first select a suitable language $\mathcal{L}$ consisting of the symbols corresponding to all relations we have the intention to consider in our structure(s). So after that those structure(s) can be seen as $\mathcal{L}$-structure(s). Then, we can use symbols in $\mathcal{L}$ as well as variable symbols (which are explained below) and logical symbols (consisting of connectives and quantifiers as explained below) to write formal logical expressions which are called by the names formulas, statements and sentences, describing our $\mathcal{L}$-structure(s) logically.

We assume that a language $\mathcal{L}$ includes a distinguished relation symbol $e$ with arity two for representing the equality relation. As will be explained more in Definition 2.3, logical symbols consist of the binary functions $+, \cdot$, and $\mid \mid$ which is the unary absolute value function and also a $0$-ary function $r$ (for every real number $r$). These functions are called and considered as connectives. The integration symbol $\int$ is also a logical symbol which functions as a quantifier. We also use an infinite list $x, y, ...$ as symbols for individual variables. We call the family of all variable and constant symbols described above, the collection of $\mathcal{L}$-terms.

**Definition 2.2.** Assume that $\mathcal{L}$ be a relational language. A simple (relational) $\mathcal{L}$-structure (or simply, a $\mathcal{L}$-structure) is a non-empty measure space $(M, \mathcal{B}, \mu)$ in which every single element is measurable (being viewed as a subsets). Moreover, $\mu(M) = 1$ and for each $n$-ary relation symbol $R \in \mathcal{L}$ (if there is any), there exists a measurable function $R^M : M^n \rightarrow \mathbb{R}$ bounded by a bound, say $\triangleright_R$. We refer to $R^M$ as the interpretations (in $M$) of the relation symbol $R$.

Note that in every structure, the binary equality relation $e(x, y)$ is interpreted as a two variable function taking value $1$ if $x = y$ and $0$ otherwise. For a language $\mathcal{L}$, the family of $\mathcal{L}$-formulas is inductively defined as follows.

**Definition 2.3.**
1. If $R$ is a $n$-ary relation symbol in $\mathcal{L}$ and $t_1, ..., t_n$ are $\mathcal{L}$-terms, then $R(t_1, ..., t_n)$ is a formula.
2. For any number $r \in \mathbb{R}$, $r$ is a formula.
3. If $\phi$ and $\psi$ are two formulas, then all $\lnot \phi$, $\phi + \psi$, $\phi \lor \psi$, $\phi \land \psi$ and $\phi \cdot \psi$ are formulas.
4. If $\phi(x, y)$ is a formula, then $\int \phi(x, y)dy$ is also a formula.

**Definition 2.4.** Free variables of formulas are easily defined (by induction) as the variables which are not bounded by the quantifiers $\int$.
Example: In the formula \( \int (2x + y) \, dy + |3z| \), the variables \( x \) and \( z \) are free variable symbols while \( y \) is bounded by the quantifier \( \int \).

We use the notation \( \phi(x_1, \ldots, x_n) \) to indicate that all free variables of the formula \( \phi \) appear in the set \( x_1, \ldots, x_n \). A closed formula is a formula which has no free variables. If \( \phi(\bar{x}) \) is a formula and \( \bar{a} \in M^{|\bar{x}|} \), the value of \( \phi(\bar{a}) \) in \( M \), is defined inductively in the natural way and is usually denoted by \( \phi^M(\bar{a}) \). For example we have

\[
(\phi + \psi)^M(\bar{a}) = \phi^M(\bar{a}) + \psi^M(\bar{a}), \\
(\phi \lor \psi)^M(\bar{a}) = \phi^M(\bar{a}) \lor \psi^M(\bar{a}),
\]

\[
\left( \int \phi(x, y)dy \right)^M(\bar{a}) = \int_M \phi^M(\bar{a}, y)dy.
\]

Therefore, \( \phi(\bar{x}) \) gives rise to a real-valued function on the domain \( M^{|\bar{x}|} \). This real-valued function is called the interpretation of the formula \( \phi \) and is denoted by \( \phi^M(\bar{a}) \). Note that, in particular, if \( \phi \) is a closed formula, then for any model \( M \), \( \phi^M(\bar{a}) \) is uniquely determined and is a real number. For instance, if \( \phi = \int \psi(y)dy \) where \( \psi(y) \) is a formula, then we have \( \phi^M = \int_M \psi^M(y)dy \). A statement is an expression with the form \( \phi(x) \geq r \) or \( \phi(x) = r \) where \( \phi(x) \) is some formula and \( r \in \mathbb{R} \). If \( \phi \) is a closed formula, then the statement is called a closed statement (or sentence).

**Definition 2.5.** Any set of closed statements is called a theory.

It is clear that expressions such as \( \phi(x) \leq r \), \( \phi(x) \geq \psi(x) + r \) or \( \phi(x) = \psi(x) + r \) (where \( \phi(x) \) and \( \psi(x) \) are formulas) are also statements since they can be written in the form \( -\phi(x) \geq -r \), \( \phi(x) - \psi(x) \geq r \) or \( \phi(x) - \psi(x) = r \) while \( -\phi(x) \) and \( \psi(x) - \phi(x) \) are again formulas. A closed statement such as \( \phi = r \) or \( \phi \geq r \) is satisfied in a \( \mathcal{L} \)-structure \( M \) (denoted by the notations \( M \models "\phi = r" \) and \( M \models "\phi \geq r" \)), if \( \phi^M = r \) and \( \phi^M \geq r \) respectively. A simple \( \mathcal{L} \)-structure \( M \) is a model of a theory \( T \) (and is denoted by \( M \models T \)), if each of the statements in \( T \) is satisfied in \( M \).

**Definition 2.6.** We call a theory satisfiable if it has a model. Also we call a theory finitely satisfiable if every finite subset of it has a model. The theory of a structure \( M \) is the family of the statements satisfied in it.

The main logical tool used in this paper is the following theorem which is called the logical compactness theorem in the literature. The interested reader can refer to for example paper [1], [2] or [3] to see more history about this result.

**Theorem 2.7.** Let \( T \) be a finitely satisfiable theory. Then it is satisfiable.

We will use the above theorem as a key ingredient in the proof of Theorem 3.6.

3. Compactness in logic and representation theorem

In this section, we use the logical compactness theorem (Theorem 2.7) to give a new logical proof for the a well-known representation theorem in analysis (see Theorem 3.6). Our proof (in below) can be viewed as an instance of how the logical techniques can be used in measure theory for proving measure existence results.

Before starting to prove the main result, in the following we mention some examples of basic measure theoretic properties expressible in the logical setting of integration logic.

**Example:** Let \( \phi(x) \) and \( \psi(x) \) be two formulas in a language \( \mathcal{L} \) with the same free variables \( x \). Then, the expressions "\( \phi(x) = 0 \) almost everywhere" and "\( \phi(x) = \psi(x) \) almost everywhere" can be stated (in the setting of the integration logic) by the two closed statements \( \int |\phi(x)| \, dx = 0 \) and \( \int |\phi(x) - \psi(x)| \, dx = 0 \) respectively.

**Example:** Let \( B = \{b_1, \ldots, b_n\} \) be any finite subset of \( \mathbb{R} \). Then, the expression "\( \psi \) gets values in \( B \) almost everywhere" is expressible by the closed statement

\[
\int \left| \left( \psi(x) - b_1 \right) \left( \psi(x) - b_2 \right) \ldots \left( \psi(x) - b_n \right) \right| dx = 0.
\]

Now we review some notions used in the main result of the paper.

**Definition 3.1.** We call a Boolean algebra \( \sigma \)-complete if every countable non-empty subset \( a_1, a_2, \ldots \) of it has a least upper bound \( \bigvee_i a_i \) (or \( \sup_{i<\omega} a_i \)) and also a greatest lower bound \( \bigwedge_i a_i \) (or \( \inf_{i<\omega} a_i \)).

105
Definition 3.2. By a measure algebra we mean a a-complete Boolean algebra $(B, \land, \lor, \neg, 0, 1)$ equipped with a map $\mu : B \to [0, \infty]$ such that the following hold.

1. $\mu(a) = 0$ if and only if $a = 0$,
2. If $a_1, a_2, \ldots$ are pairwise disjoint (i.e. $a_i \land a_j = 0$ for every distinct $i$ and $j$), then $\mu(\lor a_i) = \sum \mu(a_i)$.

Moreover, if $\mu(1) = 1$, then we call our measure algebra a probability algebra.

Note that the notations $\land$, $\lor$ and $\neg$ we are using here stand for the corresponding operations in the Boolean algebra $B$. They should not be confused with the same notations $\land$ and $\lor$ defined before which stood for the "max" and "min" of two functions.

Definition 3.3. By a $\sigma$-order-continuous isomorphism (or sequentially order-continuous isomorphism) between two measure algebras $B_1$ and $B_2$ we mean a measure preserving Boolean isomorphism $\phi : B_1 \to B_2$ such that we have $\phi(\lor a_i) = \lor \phi(a_i)$ for every increasing sequence $a_1, a_2, \ldots$ in $B_1$.

We remind that in any Boolean algebra, a natural partial order relation $\leq$ is defined by letting $a \leq b$ if and only if $a \land b = a$.

Definition 3.4. Let $(M, A, \bar{\mu})$ be a probability space. Then one can associate a probability algebra to $(M, A, \bar{\mu})$ as follows. We say that $X_1, X_2 \in A$ are equivalent if their symmetric difference is null. We denote the equivalence class of $X$ by $[X]$. Then, the family of equivalence classes forms a Boolean algebra in the natural way. Also the function $\mu([X]) = \bar{\mu}(X)$ makes of it a probability algebra.

Definition 3.5. We call a probability algebra $(B, \mu)$ probability-space-representable if there is a probability space whose associated probability algebra is $\sigma$-order-continuous isomorphic to $(B, \mu)$.

Now, we state and give a new proof for the classical representation theorem mentioned above.

Theorem 3.6. Every probability algebra $(B, \mu)$ is probability-space-representable.

Proof. Let $\mathcal{L}$ be a language consisting of a unary relation symbol $R_a$ for each $a \in B$. Let $T$ be a $\mathcal{L}$-theory consisting of the following expressions which can be written as some closed statements in integration logic in the language $\mathcal{L}$. For stating them one can get help from explanations and examples we gave earlier concerning writing mathematical properties as closed statements.

1. $R_a(x) \equiv 0$ or $1$ (for each $a \in B$),
2. $\int R_a(x) dx = \mu(a)$ (for each $a \in B$),
3. $R_{a\lor b}(x) \equiv R_a(x) \lor R_b(x)$ (for each $a, b \in B$),
4. $R_{a\lor b}(x) \equiv 1 - R_a(x)$ (for each $a \in B$).

We remind that there is difference between the meaning of the notation $\lor$ in the left and right sides of the equality in axiom 3. Indeed, in the left side it addresses the Boolean algebra operation while in the right side it stands for the logical connective "max" between two formulas. We will show that $T$ is finitely satisfiable (see Definition 2.6). Let $T_0 \subseteq T$ be a finite subset of axioms. Let $B_0$ be a finite sub measure algebra of $B$ containing those $a \in B$ such that $R_a$ appears in at least one of the axioms in $T_0$. Also let $M := \{a_1, \ldots, a_k\}$ be the atomic elements of $B_0$ (where we call $a \in B_0$ an atom of $B_0$ if for any $b \in B_0$ with $b \leq a$, we have either $b = 0$ or $b = a$). Then, it is not hard to verify that $\mu$ induces a probability measure, say $\nu$, on the finite space $(M, P(M))$. We show the satisfiability of $T_0$ by constructing a model of it over the underlying finite measure space $M = (M, P(M), \nu)$. For that, we need to interpret all relation symbols $R_a$’s $(a \in B)$ in $M$. For each $a \in B_0$, interpret the relation symbol $R_a$ with the function $R_a^M$ defined by $R_a^M(a_i) = 1$ if $a_i \leq a$ and $R_a^M(a_i) = 0$ otherwise, for any $a_i \in M$. Also for any $a \in B \setminus B_0$, we interpret the relation symbol $R_a$ with any arbitrary $\{0, 1\}$-valued function on set $M$. Now, after this interpretation it is not very hard to observe that the resulting $\mathcal{L}$-structure is indeed a model of $T_0$. It shows that $T$ is a finitely satisfiable theory. Now we use logical compactness theorem (Theorem 2.7). By this theorem, $T$ has a model where we denote by $(M, C, \bar{\mu}; R_a^M)_{a \in B}$, where each $R_a^M$ is the interpretation of the relation symbol $R_a$ in this model. Note that by the definition of a model we have $\bar{\mu}(M) = 1$ which means that $\bar{\mu}$ is a probability measure on $M$. Also each $R_a^M$ is a measurable function on $M$ with respect to the $\sigma$-algebra $C$. Let $\mathcal{B} \subseteq \mathcal{C}$ be the smallest $\sigma$-algebra making every function $R_a^M$ measurable. Also restrict $\bar{\mu}$ to $\mathcal{B}$ and for simplicity still denote the restricted measure by the same notation $\bar{\mu}$. We claim that $(M, \mathcal{B}, \bar{\mu})$ is a measure space whose associated probability algebra is $\sigma$-order-continuous isomorphic to $B$ and is in fact the desired measure space we were looking for.
It is not very difficult to observe that each $R_0^M$ is equal to a characteristic function up to a null set (use axiom 1 for showing that). Let the set $X_a$ to be defined by $\{x \in M : R_0^M(x) = 1\}$ for every $a \in B$. It is clear that every $X_a$ belongs to $\mathcal{B}$. Define $D$ to be the probability algebra associated to $(M, \mathcal{B}, \mu)$. Also, for every $A \in \mathcal{B}$, let $[A]$ to be the equivalence class of $A$ in the probability algebra $D$. Since each $R_0^M$ is a characteristic function, of course up to a null set, of the subset $X_a$, it is easy to see that the measure algebra $D$ is the same as the measure algebra associated to the restriction of the measure space $(M, \mathcal{B}, \mu)$ to the sub-$\sigma$-algebra generated by $X_a$’s. We define the function $\phi : B \rightarrow D$ by letting $\phi(a)$ to be $[X_a]$. We claim that the function $\phi$ is a measure algebra $\sigma$-order-continuous isomorphism. We first check that $\phi$ is injective. Take some $a, b \in B$ and assume that $\phi(a) = \phi(b)$. Hence, we have $[X_a] = [X_b]$. It follows that $X_a$ and $X_b$ are almost-everywhere (with respect to the measure $\mu$) equal. Then $X_a \Delta X_b$ is null. Define $a \Delta b := (a \wedge b') \lor (a' \land b)$. Obviously, $a \Delta b$ is in $B$. A suitable use of the axioms implies that $X_a \Delta X_b \triangleq a \Delta b$. Thus, $\mu(X_a \Delta b) = 0$. So, by using axiom 2, we have $\mu(a \Delta b) = \int R_0^M \, d\mu$. On the other hand, we have $\int R_0^M \, d\mu = \mu(X_a \Delta b) = 0$. Therefore, $\mu(a \Delta b) = 0$. Now, by using the definition of a probability algebra it is followed that $a \Delta b = 0$. That in turn implies that $a = b$. It completes the proof of injectivity of $\phi$. It is also straightforward to see that for every $a \in B$, $\phi(a') = \phi(a)'$.

Now we claim that if $(b_i)_{i<\omega}$ is a sequence of elements of $B$, then $\phi(\bigvee_{i<\omega} b_i) = \bigvee_{i<\omega} \phi(b_i)$ and $\phi(\bigwedge_{i<\omega} b_i) = \bigwedge_{i<\omega} \phi(b_i)$. We start to prove this. First assume that $(b_i)_{i<\omega}$ is an increasing sequence of elements of $B$ and let $b := \sup_{i<\omega} b_i$. By suitable using of axioms, it is straightforward to see that for each $i$ we have $X_{b_i} \subseteq X_b$ and $X_b \triangleq X_{b_{i+1}}$. Therefore, we have $\bigcup_{i<\omega} X_{b_i} \subseteq X_b$. On the other hand, again by using axioms, we have $\mu(X_b) = \int R_0^M = \mu(b)$ and similarly, $\mu(X_{b_i}) = \mu(b_i)$ for each $i$. Since $(b_i)_{i<\omega}$ is an increasing sequence in the measure algebra $B$, it is not hard to see that we have $\mu(\sup_{i<\omega} b_i) = \lim_{i \rightarrow \infty} \mu(b_i) = \sup_{i<\omega} \mu(b_i)$. So we get $\mu(\bigcup_{i<\omega} X_{b_i}) = \sup_{i<\omega} \mu(X_{b_i}) = \sup_{i<\omega} \mu(b_i) = \mu(\sup_{i<\omega} b_i) = \mu(b) = \mu(X_b)$. Putting all these together follows that $X_b \triangleq \bigcup_{i<\omega} X_{b_i}$. Thus, $[X_b] = [\bigcup_{i<\omega} X_{b_i}]$. In addition, we have

$$\phi\left(\bigvee_{i<\omega} b_i\right) = \phi(b) = [X_b] = \left[\bigcup_{i<\omega} X_{b_i}\right] = \bigvee_{i<\omega} [X_{b_i}] = \bigvee_{i<\omega} \phi(b_i). \quad (\ast)$$

Now assume that $(b_i)_{i<\omega}$ is an arbitrary sequence of members of $B$ which is not necessarily increasing. Let $b := \sup_{i<\omega} b_i$ and $c_i := \bigvee_{j=i}^{j=i} b_j$. Now $(c_i)_{i<\omega}$ is an increasing sequence. By $(\ast)$, we have $\phi(\bigvee_{i<\omega} c_i) = \bigvee_{i<\omega} \phi(c_i)$. So we have $\phi(\bigvee_{i<\omega} b_i) = \phi(\bigvee_{i<\omega} c_i) = \bigvee_{i<\omega} [X_{c_i}] = \bigvee_{i<\omega} [X_{\bigvee_{j=i}^{j=i} b_j}] = \bigvee_{i<\omega} [X_{b_i}] = \bigvee_{i<\omega} \phi(b_i)$. It follows that $\phi(\bigvee_{i<\omega} b_i) = \bigvee_{i<\omega} \phi(b_i)$. Moreover, by using this, we also have $\phi(\bigwedge_{i<\omega} b_i') = \phi(\bigvee_{i<\omega} b_i') = (\phi(\bigvee_{i<\omega} b_i'))' = (\bigvee_{i<\omega} \phi(b_i'))' = (\bigwedge_{i<\omega} \phi(b_i))' = \bigwedge_{i<\omega} \phi(b_i)$ which follows that $\phi(\bigwedge_{i<\omega} b_i) = \bigwedge_{i<\omega} \phi(b_i)$. It completes the proof of our above claim.

Now we want to give an argument for the surjectivity of $\phi$. Let $B'$ be the sub-$\sigma$-algebra generated by $X_a$’s. We remind from above that $D$ is the same as the measure algebra associated to the measure space $(M, \mathcal{B}', \bar{\mu}|_{B'})$. Note that by definition of a generated $\sigma$-algebra, $B'$ is the closure of the family of basic sets $X_a$’s under the operations countable unions, countable intersections and complement. So, it is not hard to verify that if we show that for any sequence $(b_i)_{i<\omega}$ of elements of $B$, $\bigvee_{i<\omega} [X_{b_i}]$ and $\bigwedge_{i<\omega} [X_{b_i}]$ are in the image of $\phi$, then it would be sufficient for concluding the surjectivity of $\phi$. But by the claim mentioned above, we have $\bigvee_{i<\omega} [X_{b_i}] = \bigvee_{i<\omega} \phi(b_i) = \phi(\bigvee_{i<\omega} b_i) \in \phi(B)$ and $\bigwedge_{i<\omega} [X_{b_i}] = \bigwedge_{i<\omega} \phi(b_i) = \phi(\bigwedge_{i<\omega} b_i) \in \phi(B)$. It follows the surjectivity of $\phi$. In a similar way, using the above claim and arguments, it is not difficult to see that $\phi$ is furthermore a $\sigma$-order-continuous and measure-preserving Boolean isomorphism. Therefore, it is a measure algebra $\sigma$-order-continuous isomorphism. It completes the proof.

Acknowledgement. The author is indebted to Institute for Research in Fundamental Sciences, IPM, for support. This research was in part supported by a grant from IPM (No. 99030116)

References


Please cite this article using:
DOI: 10.22060/ajmc.2021.19318.1045