SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Topoi From Quantum Theory and Their Logic

av

Viktor Henriksson

2018 - No M8

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

Topoi From Quantum Theory and Their Logic

Viktor Henriksson

Självständigt arbete i matematik 30 högskolepoäng, avancerad nivå

Handledare: Peter Lumsdaine

2018

Topoi From Quantum Theory and Their Logic

Viktor Henriksson

August 29, 2018

Abstract

There are two major approaches for using topoi for the interpretation of quantum mechanics, the contravariant and the covariant approach pioneered by Isham–Butterfield [1–4] and Landsman–Spitters–Heunen [5] respectively. To each topoi, there is an internal language in which constructive logic can be interpreted. This internal language is heavily used in the covariant approach, whereas any use in the contravariant approach is much more hidden.

There is also a related approach, used to study more general forms of measurement scenarios as sheaves, pioneered by Abramsky–Brandenburger [6]. This approach is general enough to make the contravariant models a special case.

In this thesis, we develop a method for considering (a variation of) these general measurement scenarios in a covariant setting using some of the tools developed by Landsman–Spitters–Heunen. The measurement scenarios are first presented as contravariant functors on a poset. We then topologize the poset using the Alexandrov topology and look at the presheaf as a bundle. This gives us a locale in the category of covariant presheaves.

We show that the locales constructed in this way are always internally spatial. We also use the constructive Gelfand duality described by Banaschewski–Mulvey [7] to get an internal commutative C^*-algebra corresponding to this locale. Lastly, we look at how to construct a state as defined by Landsmann–Spitters–Heunen.

We also consider two examples thoroughly, Spekken's Toy Model [8] and the Popescu–Rohrlich-box [9]. In doing this, we construct a new proof that the Popescu–Rohrlich-box cannot be modelled by quantum theory, without using the Tsirelson's bound.

Acknowledgements

I would first like to thank my supervisor Dr Peter LeFanu Lumsdaine for all his help with this thesis. From the first seeds to the finished work, I have felt great support. I have been free to develop my own ideas, but also to get help when needed.

Next, I would like to thank my referee Prof Dr Erik Palmgren for his quick and useful feedback.

I would also like to thank the rest of my friends and colleagues at the Department of Mathematics at Stockholm University for all their help and encouragement throughout my studies. In particular, I would like to extend my sincerest gratitude to the rest of the amanuensis for all those times in front of the whiteboard, making mathematics just as lively and creative as it is supposed to be.

Finally, to all my family and friends outside of the department; Thank you for your never-ending support. I am grateful for the way you have always let me follow my passion for physics and mathematics, but also for making me think about something else from time to time. Without you, none of this would have been possible.

Contents

Chapter 1

Introduction

From its very beginnings, the interpretation of Quantum Mechanics has been somewhat of a mystery, even among many of its great pioneers. This is perhaps best captured in the quote by Feynman:

> "I think I can safely say that nobody understands Quantum Mechanics".

One difficulty has been the non-locality of Quantum Mechanics, i.e. the fact that measurements seem to correlate over a distance without the possibility to communicate. One possible solution to this would be to assume that Quantum Mechanics was incomplete, i.e. that there are some so-called hidden variables which are yet to be discovered. This argument was perhaps most famously presented by Einstein, Podolsky and Rosen [10]. However, it was later proved by Bell [11] that any such variables needed to have a global character.

Since then, these phenomena have been explored further. It turns out that non-locality can be seen as a special case of a more general concept called contextuality. For a system to be contextual means briefly that there is no way to assign values to all possible outcomes of a measurement in a consistent way, something we will give a more detailed description of in section 2.2. That quantum systems are contextual was shown by Kochen and Specker [12].

During the last 20 years, a new approach to the interpretation of Quantum Mechanics has been explored, inspired by a series of articles by Isham and Butterfield [1–4]. In this approach, a quantum system is modelled by a presheaf, and contextuality can be formulated by properties of this presheaf. A related approach of the same flavour, heavily inspired by the first, has been explored by Landsmann, Spitters and Heunen [5]. These are the so-called topos theoretic approaches to quantum mechanics.

One approach which is often taken when discussing quantum phenomena like contextuality is to talk about toy models. These are models which share some features with quantum theory but differs in some other. Some well known such models are Spekkens toy model [8] and the Popescu-Rohrlich box [9] which will be discussed in some detail later.

In 2011, Abramsky & Brandenburger presented a sheaf-theoretic model for the description of general, possibly contextual, systems. Thus, in particular, it provides a ground for talking about the toy models of quantum mechanics, which usually has rather different origins, in a common language.

The models created by Isham–Butterfield, Landsmann–Spitters–Heunen and Abramsky–Brandenburger all give rise to presheaves and sheaves. Presheaves and sheaves can be placed in "worlds" different from the world of sets in which most of classical mathematics takes place. These worlds do not in general obey the laws of classical logic; this has been mentioned as one of the main reasons to study Topos Quantum Mechanics [13].

However, in the contravariant approach, the use of the topos-theoretic internal language is somewhat hidden. In [14] a comparison of the two approaches is made, and the author states that it is not clear to what extent the internal language is important in the contravariant approach.

In the Landsman–Spitters–Heunen approach, the place of the internal language is much more clear. In fact, many important concepts are defined in this language.

The aim of this thesis is to combine the toy model approach with the Landsmann–Spitters-Heunen approach. That is, to set up a framework for studying general measurement scenarios with the tools developed by Landsmann–Spitters–Heunen and others, and study these scenarios using this framework.

This thesis will assume familiarity with category theory, in particular limits, adjoints, and the Yoneda lemma. For the reader unfamiliar with these concepts, we refer to the textbook of Awodey [15]. A short introduction aimed towards topos theory can also be found in the Chapter Categorical Preliminaries of MacLane & Moerdijk [16].

Chapter 2

Quantum Mechanics

2.1 States and observables

In order to describe some physical system using quantum mechanics, there are three main things we would like to consider. The current state of a system, values of different so-called observables of the system, and the time evolution of the system. An observable is some physical property of the system that can be measured. For instance, for a classical particle moving in space, its velocity would be an observable.

Mathematically, quantum theory is the theory of Hilbert spaces. Indeed, we associate to each real-world quantum system a Hilbert space ([17] §3.6.1). We will however not care too much about the connection between the real world and the mathematical theory. Instead, we will just work purely on the mathematical side. However, the names of many concepts will have names inspired by what they are meant to represent. A first such example is that of a state.

Definition 2.1 ([17] §3.6.1)**.** Given a Hilbert space $\mathcal{H}$, A *(pure) state* of $\mathcal{H}$ is a unit vector in $\mathcal{H}$,

In order to use the theory to model the real world, one would first need to decide what Hilbert Space that is suitable for modelling the system of interest. In the later discussion, we will often consider the least complicated cases of $\mathcal{H} = \mathbb{C}^n$ for some n. It should be noted that these Hilbert Spaces also has real-world applications, for instance in the modelling of polarization of light or the spin of elementary particles ([18] §1.3).

We will adopt the Dirac notation of writing a state as $|x\rangle$ where x is a chosen name for the state. If a state $|x\rangle$ is a linear combination of some states, say $|0\rangle, |1\rangle$, this is written as

$$|x\rangle = a_0|0\rangle + a_1|1\rangle. \tag{2.1}$$

By Riesz representation theorem ([19] 6.2.4), it is known that each element f in the dual space of $\mathcal{H}$ can be represented by a vector $|y\rangle$ in $\mathcal{H}$ by setting $f(-) = - \cdot |y\rangle$. Here $\cdot$ denotes the inner product. Suggestively we write $f = \langle y|$, and $f(|z\rangle) = \langle y|z\rangle$. Since $\langle y|$ is linear, this also allows us to write things like

$$\langle y|x\rangle = \langle y|(a_0|0\rangle + a_1|1\rangle) = a_0\langle y|0\rangle + a_1\langle y|1\rangle. \tag{2.2}$$

Definition 2.2. An *observable* is a self-adjoint (also called Hermitian) Operator on $\mathcal{H}$.

For a self-adjoint operator H, it is possible to write $H|x\rangle$ which would give a new vector, or $\langle x|H$ which would give a new functional. Since the operator is self-adjoint, it is unambiguous to write $\langle x|H|y\rangle$.

We interpret an observable as something on our quantum system that can be measured. If $|i\rangle$ is an eigenvector of H, the eigenvalue λ_i is supposed to correspond to the value of H when measuring $|i\rangle$. By ([20] 8.10 10.18) we know that any state can be written as a linear combination of eigenvectors of H. When doing a measurement of H on $|x\rangle$, the result of the measurement will be λ_i for some i even if $|x\rangle$ is not an eigenvector of H. The probability of obtaining λ_i when doing the measurement is given by $|\langle i|x\rangle|^2$. The expectation value of H for the state $|x\rangle$ thus becomes

$$E[H] = \sum \lambda_i \langle x|i\rangle\langle i|x\rangle = \sum \lambda_i \langle x|i\rangle |x\rangle \cdot |i\rangle = \sum \langle x|i\rangle |x\rangle \cdot \lambda_i |i\rangle = \\ = \sum \langle x|i\rangle |x\rangle \cdot H|i\rangle = \sum \langle x|i\rangle\langle i|H|x\rangle = \langle x|H|x\rangle \quad (2.3)$$

where it was used that $\sum |i\rangle\langle i|$ is just the identity matrix.

After a measurement, the state will have collapsed to a new state, one which is an eigenstate corresponding to the measured eigenvalue. This means that in general the result of a measurement can depend on the order of the measurements.

Example 2.3. Let $\mathcal{H} = \mathbb{C}^2$ and let $|0\rangle, |1\rangle$ represent the standard basis. Let

$$S_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, S_y = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, S_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.4)$$

and let $|\uparrow\rangle, |\downarrow\rangle$ be the eigenvectors of S_x, i.e. $|\uparrow\rangle = .\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), |\downarrow\rangle = .\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. Doing a measurement of S_z on $|0\rangle$ would yield 1 with certainty, and a subsequent measurement of S_x would yield 1 or -1 with equal probability. Instead first measuring S_x would still give the same probabilities for the values of S_x. However, after that measurement the state of the system will now be $|\uparrow\rangle$ or $|\downarrow\rangle$ and a subsequent measurement of S_z would also have equal probabilities for the values 1 and -1.

We will show that for commuting observables, the above order dependence of measurements does not occur.

Lemma 2.4 ([21] 5.1)**.** *Let S be a set of commuting observables on some finite dimensional Hilbert space $\mathcal{H}$. Then there exists a basis $\mathcal{B}$ such that for each $b_i \in \mathcal{B}$ and $A_i \in S$, b_i is an eigenvector of A_i.*

Theorem 2.5. *Let A and B be commuting observables on some Hilbert space $\mathcal{H}$. Then the probability to get a certain measurement result is independent of the order of measurement between A and B. Furthermore, the final state is also independent of the order.*

Proof. Let $|\varphi\rangle$ be a state in $\mathcal{H}$. By Lemma 2.4 there exists a basis $\{|i\rangle\}$ of $\mathcal{H}$ such that each $|i\rangle$ is an eigenvector of both A and B. Thus it is possible to write

$$|\varphi\rangle = \sum \lambda_i |i\rangle. \quad (2.5)$$

After a measurement of A, the state will collapse to the set of eigenstates of A which share the eigenvalue obtained in the measurement. Thus the new state will be

$$|\varphi'\rangle = \sum c\lambda_j |j\rangle. \tag{2.6}$$

where c is a normalisation factor. The probability to obtain this result is

$$p = \sum_j |\lambda_j|^2. \tag{2.7}$$

and thus $c = p^{-1/2}$. Now, measuring B, the probability to end up in the new state

$$|\varphi''\rangle = \sum_k c'c\lambda_k |k\rangle \tag{2.8}$$

is $p' = \sum c^2 |\lambda_k|^2 = \sum_k |\lambda_k|^2/p$. Thus, the probability of getting this measurement of A and B is $pp' = \sum_k |\lambda_k|^2$ which depends only on the coefficients λ_k. Doing the same calculation taking B first gives the same result, and thus the theorem follows. □

The states defined in Definition 2.1 are called pure since we have full knowledge about the state. If this is not the case, we have what is called a mixed state. The physical intuition for these is that they are to be seen as classical probability distributions of pure states, i.e. situations in which some classical lack of information prevent us from knowing which pure state we are in.

Definition 2.6 ([22] §2.1.2)**.** Suppose that an ensemble of states in $\mathcal{H}$ are given, where the frequency of the state $|\varphi_i\rangle$ is p_i. Then the matrix

$$\rho = \sum_i |\varphi_i\rangle p_i \langle\varphi_i| \tag{2.9}$$

is used to represent the quantum state of the system. It is called the *density operator*. Note that $p_j = 1$ for some j if and only if $\rho^2 = \rho$. Thus a state satisfying this is called a *pure state*. A state for which $\rho^2 \neq \rho$ is called *mixed*.

Remark 2.7. For any pure state, there is always an observable for which there is no uncertainty of the outcome, namely the projection onto the space spanned by $|\varphi\rangle$.

Since Quantum Theory takes place in vector spaces, there is a concept of tensor products of quantum states. Thus, given two Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$ with states $|a\rangle \in \mathcal{H}_1, |b\rangle \in \mathcal{H}_2$, we write things like $|a\rangle \otimes |b\rangle$. We note that if A and B are observables on $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, and has a and b as eigenvalues, then $A \otimes B$ has ab as an eigenvalue.

2.2 Contextuality

A concept that has recently received much attention, is that of quantum contextuality. In particular, it has been suggested to provide some of the power of quantum computation [23]. It is also a generalisation of non-locality, a quantum feature which has fascinated many quantum scientists.

We will use Spekkens formalization of contextuality, for which we first need to introduce the concepts of operational theories and ontological models. Intuitively, operational theories are just a list of probabilities for getting a certain outcome when doing some preparation and some measurement. It should be noted that "preparation" can be interpreted rather vaguely. For an astronomer, it could, for instance, mean to wait for a certain astronomical phenomenon to happen.

Definition 2.8 ([24] §II)**.** Let P, M and K be sets. We think of P as a set of possible preparations, M as possible measurements and K as possible outcomes. An *operational theory* is a probability mass function (or probability density function, if one allows continuous underlying spaces) $p(k, q, m)$ for each $k \in K, q \in P, m \in M$. Two preparations q, q' are called *equivalent* if $p(k, q, m) = p(k, q', m)$ for all k and m. We then write $q \cong q'$. Similarly, two measurements are equivalent, $m \cong m'$, if $p(k, g, m) = p(k, g, m')$ for all k and q.

It should be noted that in Spekkens paper, transformations are also introduced alongside preparations and measurement. We will not need that here, and thus it has been omitted. We also mention a remark of Spekkens that sometimes a transformation can be seen as a part of the preparation, which gives further motivation for our choice to omit it.

Note that the above only gives information about relations between preparations and measurements, but says nothing about what is being prepared. Although everything that we have control over is the preparation and the measurement, a theory of physics usually seek to explain what happens in between the preparation and measurement. For instance, if we prepare our system by placing a ball in an automatic release system, and then measure the sound of the ball hitting the ground, most physicists would have a rather clear idea of what state the ball was in between these two events. This motivates the following definition.

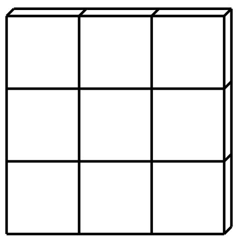
Definition 2.9 ([24] §II)**.** Let an operational theory (K, P, M, p) be given. An *ontological model* for this theory is a space Ω and probability mass (or density) functions $\mu_q : \Omega \to \mathbb{R}, \xi_{k,m} : \Omega \to \mathbb{R}$ such that

$$p(k, q, m) = \int_\Omega \mu_q(\lambda)\xi_{k,m}(\lambda)d\lambda. \tag{2.10}$$

Definition 2.10 ([24] §II)**.** An ontological model is called *preparation noncontextual* if for $q \cong q'$, $\mu_q = \mu_{q'}$. Similarly, it is called *measurement noncontextual* if for $m \cong m'$, $\xi_{k,m} = \xi_{k,m'}$. It is *(universally) noncontextual* if it is both preparation noncontextual and measurement noncontextual. An operational theory for which there is no noncontextual ontological model is called *contextual*.

We will now prove that if quantum theory is given by an ontological model, then quantum theory is contextual in the sense of Spekkens. This will be done using the Mermin–Peres magic square, presented in e.g. [25],[26].

Definition 2.11. The *Mermin–Peres magic square* is a square



together with the following game. We start with some preparation process which produces a physical state.[1] Next, we are to set up two sets of measurements, call them A and B. Each set consists of one measurement for each box in the square. These measurements must be set up with the following rules:[2]

(i) Each measurement give either -1 or 1.

(ii) The probability of getting 1 or -1 given the preparation process is known.[3]

(iii) When provided with a row and a column, we choose all measurements in the corresponding row from A, and all measurements from the corresponding column from B. We should then be able to get the following result:

   a) The outcome agree on the square shared by the row and the column.

   b) The row outcomes multiply to -1.

   c) Has the column outcomes multiply to 1.

If such a physical state is prepared, we say that the physical model *wins the magic square game.*

Lemma 2.12. *No noncontextual ontological model can win the magic square game.*

Proof. Let $p_{r,i,j}, p_{c,i,j}$ denote the probability of getting 1 in the square in row i and column j doing a row and column measurement respectively. Since these measurements need to coincide, these probabilities need to be the same (and correlated). Thus we can drop the row or column specification and talk about the probabilities $p_{i,j}$.

We note first that there is no loss of generality to consider a model where the preparation process prepares a single state λ with probability 1. Indeed, if this preparation cannot win the game, then for any other preparation, we can choose a λ which is prepared with a nonzero probability, and we are back in the first case, giving a nonzero probability to lose the game.[4]

Now the noncontextuality assumption states that these probabilities can only depend on the prepared state λ. This means that it is possible to talk about the probabilities $p_{i,j}$ also for the rows and columns that are not measured (since the state is already prepared when the choice of rows and columns are made). For each i, j, let $p'_{i,j} = p_{i,j}$ if $p_{i,j} \neq 0$. Otherwise, let $p'_{i,j} = 1$ (which is then the probability of getting -1 at the position i, j). Clearly, $\prod_{i,j} p'_{i,j} \neq 0$. Thus, there is a nonzero probability that the game would be won using a labeling on the positions where the position is labeled 1 wherever $p'_{i,j} = p_{i,j}$ and -1 otherwise. We will now show that no such labeling is possible.

Let the labels be given by:

[1] I.e a point in Ω if we assume that we are working with an ontological model.

[2] This can be phrased as a correlation game, where two players does measurements on a predetermined state. One player gets to know the row, and the other gets to know the column, and their task is to fullfill the conditions of the game. See [27].

[3] In ontological model language, this is saying that $p(1, q, m_i), p(-1, q, m_i)$ is known for $m_i \in A \cup B$.

[4] We have to be slightly more careful in the case of a continuous space Ω, and a probability distribution $\mu : \Omega \to \mathbb{R}$. However, in principle, the same argument works. Indeed, if we know that the probability of winning the game (call it $w : \Omega \to \mathbb{R}$) is less than 1 for each $\lambda \in \Omega$, then we have that the probability of winning the game is $\int_\Omega \mu(\lambda) w(\lambda) d\lambda < \int_\Omega \mu(\lambda) d\lambda = 1$.

a	b	c
d	e	f
g	h	i

.

In order to satisfy the game conditions, we would need to assign the value 1 or −1 to the variables in such a way that

$$abc = -1 \qquad def = -1 \qquad ghi = -1 \tag{2.11}$$
$$adg = 1 \qquad beh = 1 \qquad cfi = 1. \tag{2.12}$$

Multiplying these equations together, we get

$$abcdefghi = 1 \qquad abcdefghi = -1, \tag{2.13}$$

a clear contradiction. □

Lemma 2.13. *Quantum theory wins the Mermin–Peres magic square game.*

Proof. Prepare the following state

$$|\varphi\rangle = \frac{1}{2}\left((|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle) \otimes (|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle)\right). \tag{2.14}$$

Given this state, do the following measurements in the magic square

$S_x \otimes I$	$-S_x \otimes S_z$	$I \otimes S_z$
$S_x \otimes S_x$	$-S_y \otimes S_y$	$S_z \otimes S_z$
$I \otimes S_x$	$-S_z \otimes S_x$	$S_z \otimes I$

,

where $S_a \otimes S_b$ is an abbreviation of $S_a \otimes I \otimes S_b \otimes I$ when doing row measurements, and an abbreviation for $I \otimes S_a \otimes I \otimes S_b$ when doing column measurements column.

The lemma now follows by checking that the measurements in each row commutes and the measurements in each column commutes, and by checking the possible measurement outcomes. □

Theorem 2.14. *Quantum theory is contextual.*

Proof. This is an obvious corrollary of Lemma 2.12 and 2.13. □

One way to interpret this result is that quantum theory in some sense needs to be aware of what other measurements are going to be made, in this example what row and what column is chosen, before its values are specified. This choice of measurements that are simultaneously being made is what we will call a *context*.

Further discussion of contextuality will be given Chapter 5, where it will be discussed in the topos-approach sense. Further reading on non-topos approaches to contextuality can be found for instance in [12] and [28].

Chapter 3

Logic for Topoi

3.1 Many Sorted Logic

The logic we will use will generally be many sorted. A sort is some collection of elements, usually thought of as being of the same kind (or, of the same sort). When doing logic in the category **Sets**, we interpret each sort as a set. Given objects of some sort, it is possible to construct new objects, e.g. tuples. These objects will be said to have a type, described in terms of the original sort. Thus, sorts are a special kind of type, seen as a base type.

Presented with these concepts, we can begin a formal introduction of many-sorted logic.

Definition 3.1 ([29] D.1.1.1)**.** A *first order signature* consists of:

(i) A set of *sorts*.

(ii) A set of *function symbols*, and for each function symbol, a tuple of sorts representing its type. For f of type $A_1, \cdots, A_n, B$, we write $f : A_1 \cdots A_n \to B$ where the A_i can be thought of as source types, and B as a target type. The number n is the *arity* of f. A function symbol of arity 0 is also called a constant.

(iii) A set of *relation symbols*, and for each relation symbol, information about its type. For R of type $A_1, \cdots, A_n$, write $R \rightarrowtail A_1 \cdots A_n$. Again, n is the *arity* of R.

Definition 3.2 ([29] D.1.1.2)**.** For a signature Σ, the collection of *terms* over Σ is defined recursively by the following rules:

(i) For each sort A, a variable $x : A$ is a term of type A.

(ii) For each function symbol $f : A_1 \cdots A_n \to B$, if $t_1 : A_1, \cdots, t_n : A_n$, then $f(t_1, \cdots, t_n)$ is a term of type B.

We note that variables are defined as being of some certain sort. Thus, there is no such thing as a variable without a sort.

Definition 3.3 ([29] D.1.1.3)**.** For a signature Σ, the collection of *first-order formulae* over Σ is defined recursively by the rules below. The set of *free variables* of a formula is defined simultaneously.

(i) For any relation symbol R, and terms $t_1, \cdots, t_n$, $R(t_1, \cdots, t_n)$ is a formula with free variables the union of the free variables in the t_i. For each type A, a special relation symbol $=_A$ of type A, A is always required.

(ii) Truth $\top$ is a formula with no free variables.

(iii) Falsity $\bot$ is a formula with no free variables.

(iv) For any two formulas φ, ψ, the binary conjunction $\varphi \wedge \psi$, the binary disjunction $\varphi \vee \psi$ and implication $\varphi \Rightarrow \psi$ are formulae. Their free variables are the union of the free variables in φ and ψ. A particular case of implication, $\varphi \Rightarrow \bot$ is called negation and denoted $\neg\varphi$.

(v) For any formula φ with free variable x, existential quantification $\exists x\varphi$ and universal quantification $\forall x\varphi$ are formulae. It has as free variables the free variables of φ, except x.

Definition 3.4 ([29] D.1.1.3)**.** For a signature Σ, the class of geometric formulae is defined as for first-order formulae, with the difference that universal quantification and implication are not allowed, while infinitary disjunction, written $\bigvee_{t\in T} t$ is. Here, T is some arbitrary set of terms, with the restriction that the number of free variables is finite.

Definition 3.5 ([29] D.1.1.4)**.** A *context* is a finite list $x_1, \cdots, x_n$ of variables. If $x_i : A_i$, the list $A_1, \cdots, A_n$ is the type of the context. Given a formula φ, a context is called *suitable* for φ if all free variables of φ is in the list.

Definition 3.6. If $\vec{x}$ is a context of type $A_1, \ldots, A_n$, and $\vec{s} = s_1 \ldots s_n$ is a list of terms where s_i has type A_i, then $\varphi[\vec{s}/\vec{x}]$ denotes the formula φ where each free occurence of x_i has been replaced with s_i.

Definition 3.7 ([29] D.1.1.5)**.** Given a signature Σ, two formulas φ, ψ and a context $\vec{x}$ suitable for both of them, a *sequent (over Σ)* is an expression $\varphi \vdash_{\vec{x}} \psi$.

Definition 3.8 ([29] D.1.1.6)**.** A *theory* over Σ is a set T of sequents called the *axioms* of T. The theory is called a *first-order theory* if the axioms are first order, and a *geometric theory* if the formulae are geometric.

3.2 Heyting algebras

The internal logic of a topos, as will be introduced in Section 3.4, is generally constructive. Constructive mathematics is tightly connected to Heyting algebras. Thus, we will introduce these here.

Definition 3.9 ([16] §I.7)**.** Let x and y be elements of a poset X. A binary *meet* is an object $x \wedge y$ with the property that $x \wedge y \leq x, x \wedge y \leq y$ and if $z \leq x, z \leq x$, then $z \leq x \wedge y$. A binary *join*, denoted $x \vee y$ is defined similarly but with the inequalities reversed. A *lattice* is a poset where all elements have binary meets and joins. Note that if the poset is considered as a category, this is exactly saying that the category has binary products and coproducts.

Definition 3.10 ([16] §I.8)**.** A *Heyting algebra* is a lattice X with a smallest and largest element, denoted 0 and 1 (initial and terminal element if considered as a category), together with a binary operation $x \Rightarrow y$ characterized by the property that $z \leq (x \Rightarrow y)$ if and only if $z \wedge x \leq y$. If the lattice is considered as a category, this is exactly saying that $x \Rightarrow y$ is an exponential. Thus this operation is usually called the *exponential* of the Heyting algebra.

Proposition 3.11 ([16] §I.8)**.** *A Heyting algebra is distributive, i.e.*

$$(x \vee z) \wedge y = (x \wedge y) \vee (z \wedge y). \tag{3.1}$$

Also, the equality

$$x \to (y \wedge z) = (x \Rightarrow y) \wedge (z \Rightarrow y) \tag{3.2}$$

holds.

Proof. This is a direct consequence of looking at Heyting algebras from a categorical perspective. There is an adjunction

$$\mathrm{Hom}(z \wedge x, y) \cong \mathrm{Hom}(z, x \Rightarrow y) \tag{3.3}$$

showing that the functor $-\wedge x$ preserves coproducts (i.e. joins), and the functor $x \Rightarrow -$ preserves products (i.e. meets). □

When interpreting logic in the sense of a Heyting algebra, we think of the elements of the algebra as propositions, and the meet and join as operations corresponding to "and" and "or". The 0 and 1 then corresponds to falsity and truth, and the exponential is the implication. Talking about logic, one might also want some kind of negation. A negation in the classical sense is not possible. However, we can phrase negation in terms of implications of falsity.

Definition 3.12 ([16] §I.8)**.** Given an element x of a Heyting algebra, the *negation* or *pseudo-complement* of x is defined as $\neg x = x \Rightarrow 0$.

As we will see, the law of excluded middle, i.e. that either a proposition or its negation needs to be true, does not always hold in a Heyting algebra. We will find some comfort in the fact that both a proposition and its negation cannot be true simultaneously.

Proposition 3.13 ([16] §I.8)**.** *Given x in a Heyting algebra, $x \wedge \neg x = 0$.*

Proof. Clearly $\neg x \leq x \Rightarrow 0$ (since they are equal). By the defining property of the exponential it then follows that $\neg x \wedge x \leq 0$. □

We note that any Boolean algebra is a Heyting algebra. Indeed, in this case $x \Rightarrow y = x \vee \neg y$. However, not all Heyting algebras are Boolean, as the following example shows.

Example 3.14 ([16] §I.7)**.** Consider the set of open subsets of $[0, 1]$ ordered by inclusion. It is clear that intersections and unions define meets and joins, and that the empty set and full interval provides the 0 and the 1 respectively. Let $U, V \subseteq [0, 1]$. We can define $U \Rightarrow V$ to be the largest open W such that $W \wedge U \leq V$. We note that with this definition of the exponential, the negation of a set U is the interior of its complement. Now, let $U = [0, 1/2)$. Then $\neg U = (1/2, 0]$. Thus $U \wedge \neg U = [0, 1] \setminus 1/2$, showing that the law of excluded middle does not hold in this Heyting algebra. Since the law of excluded middle holds in all Boolean algebras, this shows that not all Heyting algebras are Boolean.

3.3 Introducing Topoi

Let us consider the definition of a topos.

Definition 3.15 ([15] 8.17)**.** A category $\mathcal{T}$ is called a *topos* if it satisfies the following properties:

- It has all finite limits,
- It has all exponentials,
- It has a subobject classifier,

where a *subobject classifier* is an object $\Omega \in \mathcal{C}$ together with an arrow $\top : 1 \to \Omega$ such that for any subobject $A \rightarrowtail X$, there is a unique arrow such that the following diagram is a pullback:

$$\begin{array}{ccc} A & \longrightarrow & 1 \\ \big\downarrow & & \big\downarrow{\scriptstyle \top} \\ X & \longrightarrow & \Omega. \end{array}$$

We note that the subobject classifier in **Sets** is the set $\mathbf{2} = \{0, 1\}$. A mapping into $\mathbf{2}$ can be described by giving a subset which is mapped to 1, and assume that everything else is mapped to 0. Thus, a map into a subobject classifier can be thought of as a question if a certain object is in a certain subset, and the subobject classifier is the object containing all possible answers to this question.

The definition of a topos is rich in consequences. Most importantly, it provides topoi as a setting for the interpretation of logic, as introduced in Section 3.4. In this section, we will consider some consequences needed for this interpretation.

Lemma 3.16 ([16] IV.5.4)**.** *A topos has all finite colimits.*

Definition 3.17 ([16] IV.6)**.** Given an arrow f, the *image* of f is a monic arrow m such that f factors through m, and such that if f factors through another monic m', then m also factors through f.

Proposition 3.18 ([16] IV.6.1)**.** *In a topos, every f has an image m, and the unique arrow e such that $f = me$ is epic.* □

Theorem 3.19 (cf. [16] IV.6.3 IV.8.1 [30] 6.2.1)**.** *Given a topos $\mathcal{T}$, and an object $A \in \mathcal{T}$. The set of subobjects of A is a Heyting algebra, where, for subobjects $T \rightarrowtail A, S \rightarrowtail A$, the operations are defined as follows.*

(i) The meet is given by the pullback

$$\begin{array}{ccc} S \cap T & \longrightarrow & T \\ \big\downarrow & & \big\downarrow \\ S & \longrightarrow & A. \end{array} \tag{3.4}$$

(ii) The join is given by the image of the unique arrow $S + T \to A$.

(iii) The implication $S \Rightarrow T$ is given by the subset specified by the equalizer of the characteristic map of $S \cap T$ and S.

(iv) The largest element 1 is $A \xrightarrow{\mathrm{id}} A$.

(v) The smallest element 0 is the unique arrow $0 \to A$. □

Theorem 3.20 ([16] IV.6.3 IV.9.3)**.** *For each arrow $k : A \to B$, the map $k^{-1} : \mathrm{Sub}(B) \to \mathrm{Sub}(A)$ defined by the following pullback square*

$$\begin{array}{ccc} k^{-1}(D) & \longrightarrow & D \\ \downarrow & & \downarrow \\ A & \longrightarrow & B \end{array} \tag{3.5}$$

is a functor, and has both a left and a right adjoint. The left adjoint is denoted $\exists_k$ *and sends* $E \xrightarrow{f} A$ *to the image of* $k \circ f$*. The right adjoint is denoted* $\forall_k$*.*

Proposition 3.21 ([16] IV.1.2)**.** *Let* f *be an arrow in a topos* $\mathcal{E}$*. If* f *is both epic and monic, then* f *is an isomorphism.*

3.4 Interpreting Logic in Topoi

Given a many-sorted theory, we might want to find models for it. One way to do this is to associate to each sort a set S. Then, any formula taking variables in this set would define a subset $U \subseteq S$. However, classically, the set of subsets of a set is a Boolean algebra. Since we want to be more general, we want to find a setting where the subobjects are not in general boolean, but rather a Heyting algebra. By Theorem 3.19, topoi is the place to look.

Definition 3.22 ([29] D.1.2.1)**.** Given a topos $\mathcal{C}$ and a signature Σ, a Σ-*structure* M in $\mathcal{C}$ is given by the following assignments:

(i) For each sort A of Σ, an object $MA \in \mathcal{C}$.

(ii) For a list of sorts $A_1, \cdots, A_n$ the object $M(A_1, \cdots, A_n) = MA_1 \times \cdots \times MA_n$.

(iii) For each function symbol $f : A_1 \cdots A_n \to B$ an arrow $Mf : M(A_1, \cdots, A_n) \to M(B)$.

(iv) For each relation $R \rightarrowtail A_1 \cdots A_n$, a subobject $MR \rightarrowtail M(A_1, \cdots, A_n)$. Note that in a topos, this could also be identified with an arrow $M(A_1, \cdots, A_n) \to \Omega$.

Definition 3.23 ([29] D.1.2.3)**.** Suppose we have a Σ-structure, and let t be a term over Σ in a suitable context $\vec{x}$. Assume further that $\vec{x}$ has type $A_1 \cdots A_n$. We write $\vec{x}.t$ to denote t as a term in the context $\vec{x}$. The interpretation of $\vec{x}.t$, denoted $\llbracket \vec{x}.t \rrbracket$ is defined recursively by the following rules:

(i) If t is a variable, then it is x_i for some x_i in $\vec{x}$. Then $\llbracket \vec{x}.t \rrbracket$ is the ith product projection. Note that in the particular case when t is $x.x$, the interpretation will just be the identity.

(ii) If t is $f(t_1, \cdots, t_m)$ where $t_i : B_i$, then arrows $\llbracket \vec{x}.t_i \rrbracket : M(A_1, \cdots, A_n) \to MB_i$ are given, and thus the arrow $(\llbracket \vec{x}.t_1, \rrbracket, \cdots, \llbracket \vec{x}.t_n \rrbracket) : M(A_1, \cdots, A_n) \to M(B_1, \cdots, B_n)$ exists. The interpretation $\llbracket \vec{x}.t \rrbracket$ is then defined to be the composition $Mf \circ (\llbracket \vec{x}.t_1, \rrbracket, \cdots, \llbracket \vec{x}.t_n \rrbracket)$.

Similarly, as above, we want interpretations of the formulae over Σ.

Definition 3.24 ([29] D.1.2.6)**.** Given a Σ-structure, a formulae in a suitable context $\vec{x}$ of type $A_1, \cdots, A_n$, will be interpreted as subobjects of $M(A_1, \cdots, A_n)$ by the following rules:

1. If φ is $R(t_1, \cdots, t_m)$ for a relation symbol R of type $B_1, \cdots, B_m$, then $[\![\vec{x}.\varphi]\!]$ is defined by the pullback

$$\begin{array}{ccc} [\![\vec{x}.\varphi]\!] & \longrightarrow & MR \\ \downarrow & & \downarrow \\ M(A_1, \cdots, A_n) & \xrightarrow{([\![\vec{x}.t_1]\!], \cdots, [\![\vec{x}.t_m]\!])} & M(B_1, \cdots, B_m) \end{array} \tag{3.6}$$

2. For $\top$, $[\![\vec{x}.\top]\!]$ is the maximal subobject of $M(A_1, \cdots, A_n)$.
3. For $\bot$, $[\![\vec{x}.\bot]\!]$ is the minimal subobject of $M(A_1, \cdots, A_n)$.
4. For $\varphi * \psi$ where $*$ is $\wedge, \vee$ or $\Rightarrow$, the interpretations are $[\![\vec{x}\varphi]\!] * [\![\vec{x}.\psi]\!]$ with the subobject operations defined in 3.19. Similarly, $[\![\vec{x}.(\neg\varphi)]\!]$ is $\neg[\![\vec{x}.\varphi]\!]$.
5. For $\exists y\varphi$ and $\forall y\varphi$, let y be of type B, and let $\pi : M(A_1, \cdots, A_n) \times M(B) \to M(A_1, \cdots, A_n)$. Then $[\![\vec{x}.\,(\exists y\varphi)]\!] = \exists_\pi [\![\vec{x}, y.\varphi]\!]$ and $[\![\vec{x}.\,(\forall y\varphi)]\!] = \forall_\pi [\![\vec{x}, y.\varphi]\!]$.

In some cases, the topos behaves nice enough that every algebra of subobjects has arbitrary joins. A topos where this is the case is called geometric. In this case we can also add the (rather obvious) interpretation:

6. For $\bigvee \varphi_i$, $[\![\bigvee \varphi_i]\!] = \bigvee_i [\![\varphi_i]\!]$.

These interpretations make it possible to interpret any (geometric or first order) formula in a topos. However, we find it of interest to be explicit about the interpretation of $\varphi[\vec{s}/\vec{x}]$.

Lemma 3.25 ([29] D.1.2.7)**.** *Let $\vec{y}$ be a context interpreted by Y, and let $\vec{s}$ be a list of terms of the same type as the variables in $\vec{y}$. Let $\vec{x}$ be a context suitable for the terms in $\vec{s}$. Then for a formula φ, the interpretation is given by the following pullback square*

$$\begin{array}{ccc} [\![\vec{x}.\varphi[\vec{s}/\vec{y}]]\!] & \longrightarrow & [\![\vec{y}.\varphi]\!] \\ \downarrow & & \downarrow \\ X & \xrightarrow{([\![\vec{x}.s_1]\!], \cdots, [\![\vec{x}.s_m]\!])} & Y \end{array} \tag{3.7}$$

□

Given these interpretations, it might be of interest to define some kind of truth. This is done in the following way.

Definition 3.26 ([29] D.1.2.12)**.** Let ς be a sequent over Σ on the form $\varphi \vdash_{\vec{x}} \psi$, and let M be a Σ-structure in a topos $\mathcal{T}$. Suppose that $\vec{x}$ has types interpreted by $X_1, \cdots, X_n$, and write $X = X_1 \times \cdots \times X_n$. We then say that M *satisfies* ς, written $M \vDash \varsigma$, if $[\![\vec{x}.\varphi]\!]_M \leq [\![\vec{x}.\psi]\!]_M$ as subobjects of X. If φ is a formula, we say that M satisfies φ, $M \vDash \varphi$, if $\top \vdash_{\vec{x}} \varphi$, i.e. if $[\![\vec{x}.\varphi]\!]_M = X$.

Definition 3.27. For a theory T over Σ, a Σ-structure M is called a *model* for T, written $M \vDash T$ if $M \vDash \varsigma$ for each $\varsigma \in T$.

A special case of Σ-structure is that corresponding to the whole topos.

Definition 3.28. Let $\mathcal{T}$ be a topos, and let Σ be the signature consisting of a sort for each object in $\mathcal{T}$, a function symbol for each arrow in $\mathcal{T}$, and a relation symbol for each monomorphism. This is called the *Mitchell–Bénabou language* of $\mathcal{T}$. It has a rather obvious interpretation $M = \mathcal{T}$. Thus we write can also write $\mathcal{T} \vDash \varsigma$ for a sequent (or formula) ς.

3.5 Constructive reasoning

In classical logic, we are used to a number of rules for reasoning about logical formulae. For instance, we know that if we know φ and ψ, then we know $\varphi \wedge \psi$. This is usually formalized by the rule

$$\frac{\varphi \qquad \psi}{\varphi \wedge \psi}\,. \tag{3.8}$$

We will now introduce the rules which we may use when reasoning constructively. We will use the abbreviation that

$$\frac{\varphi}{\overline{\psi}} \quad \text{is equivalent to} \quad \frac{\varphi}{\psi} \quad \text{and} \quad \frac{\psi}{\varphi}\,. \tag{3.9}$$

The rules of inference for constructive logic are the following [29].

1. The *identity axiom* $\varphi \vdash_{\vec{x}} \varphi$,
2. The *substitution rule*,
$$\frac{\varphi \vdash_{\vec{x}} \psi}{\varphi[\vec{s}/\vec{x}] \vdash_{\vec{x}'} \psi[\vec{s}/\vec{x}]} \tag{3.10}$$
where $\vec{x}'$ is such that it contains all variables in $\vec{s}$.
3. The *cut rule*
$$\frac{\varphi \vdash_{\vec{x}} \varsigma \qquad \varsigma \vdash_{\vec{x}} \psi}{\varphi \vdash_{\vec{x}} \psi}\,. \tag{3.11}$$
4. The *equality axioms* $(\top \vdash_{\vec{x}} (x = x)$ and $((\vec{x} = \vec{y}) \wedge \varphi) \vdash_{\vec{x}} \varphi[\vec{y}/\vec{x}]$.
5. The *finite conjuction rules* $\varphi \vdash_{\vec{x}} \top$, $(\varphi \wedge \psi) \vdash_{\vec{x}} \varphi$, $(\varphi \wedge \psi) \vdash_{\vec{x}} \psi$, and
$$\frac{\varsigma \vdash_{\vec{x}} \varphi \qquad \varsigma \vdash_{\vec{x}} \psi}{\varsigma \vdash_{\vec{x}} (\varphi \wedge \psi)}\,. \tag{3.12}$$
6. The *finite disjunction rules* $\perp \vdash_{\vec{x}} \varphi$, $\varphi \vdash_{\vec{x}} (\varphi \vee \psi)$, $\psi \vdash_{\vec{x}} (\varphi \vee \psi)$, and
$$\frac{\varphi \vdash_{\vec{x}} \varsigma \qquad \psi \vdash_{\vec{x}} \varsigma}{(\varphi \vee \psi) \vdash_{\vec{x}} \varsigma}\,. \tag{3.13}$$
7. The *implication rule*
$$\frac{(\varphi \wedge \psi) \vdash_{\vec{x}} \varsigma}{\overline{\psi \vdash_{\vec{x}} (\varphi \Rightarrow \varsigma)}}\,. \tag{3.14}$$
8. The *exsistensial quantification rule*
$$\frac{\varphi \vdash_{\vec{x}} \psi}{\overline{(\exists y)\varphi \vdash_{\vec{x}} \psi}}\,. \tag{3.15}$$

For first order logic we also have

9. The *universal quantification rule*

$$\frac{\varphi \vdash_{\vec{x}.y} \psi}{\varphi \vdash_{\vec{x}} (\forall \psi)} . \tag{3.16}$$

and for geometric logic we have

10. The *Frobenius axiom* $\varphi \wedge (\exists y)\psi \vdash_{\vec{x}} (\exists y)(\varphi \wedge \psi)$
11. The *infinitary conjunction and disjunction rules* which are the infinte versions of (5) and (6).
12. The infinitary distribution axiom $\varphi \wedge \bigvee \psi_i \vdash_{\vec{x}} \bigvee(\varphi \wedge \psi_i)$.

Given rules of inference and some rules for their interpretation, we want to know that this correspondence is sound. That is, we want to know that any model which satisfy the axioms also satisfy the conclusions.

Theorem 3.29 ([29] D.1.3.2)**.** *Let T be a (first order or geometric) theory over a signature Σ. If M is a model of T in a topos, and if ς is provable using the appropriate rules of inference, then $M \vDash \varsigma$. Note that the topos needs to be geometric for this to make sense for a geometric theory.* □

Chapter 4

Presheaves, Sheaves & Locales

4.1 Presheaf Topoi

One important topos is that of presheaves, i.e. functors from a small category into **Sets**. Here, small means that the collection of objects and arrows in the category is a set. These topoi have some nice properties that we will explore in this section.

Definition 4.1 ([16] p. 37)**.** Given a category $\mathcal{C}$ and an object $C \in \mathcal{C}$, a set s of arrows with codomain C is called a *sieve* if $f \in s$ implies that $f \circ g \in s$ for all g with codomain the domain of f. Given an arrow $h : D \to C$, it is clear that the set $\uparrow h = \{h \circ h' | h' \text{ has codomain } D\}$ is a sieve.

As an example of the above definition, let us consider the case when the category is a poset. Then the definition says that a sieve is a downwards closed set, that is, a set S such that $s \in S$ and $t \leq s$ implies $t \in S$. In this case we also use the notation $\uparrow s = \{t \mid t \leq s\}$ since the maps are implicit.

Theorem 4.2 ([15] 8.18)**.** *Any functor category* $\mathbf{Sets}^{\mathcal{C}^{op}}$ *where* $\mathcal{C}$ *is small is a topos. The subobject classifier is given by the functor defined by sending* $C \in \mathcal{C}$ *to the set of sieves on* C*, and sending* $f : D \to C$ *to the map* $\Omega(f)(S) = \{t \mid f \circ t \in S\}$*. The arrow* $1 \xrightarrow{\top} \Omega$ *is given by mapping the single element in* $1(C)$ *to* $\uparrow \mathrm{id}_C$*.* □

The following result is useful when wanting to do explicit calculations.

Theorem 4.3 ([15] 9.17)**.** *For* $f : \mathcal{C} \to \mathcal{D}$ *a functor between small categories, the functor* $f^* : \mathbf{Sets}^{\mathcal{D}^{op}} \to \mathbf{Sets}^{\mathcal{C}^{op}}$ *defined by*

$$f^*(F)(C) = F(fC) \tag{4.1}$$

has both a left and right adjoint. □

Consider specifically the case when $\mathcal{C}$ is the one object category with only the trivial arrow. Then $\mathbf{Sets}^{\mathcal{C}^{op}} \cong \mathbf{Sets}$. For each $D \in \mathcal{D}$, let f_D be the functor mapping the single element to the element D. Then f_D^* is just evaluation at D. By the above, this functor has (and thus is itself) both a left and a right adjoint. This means that the evaluation functor preserves both limits and colimits. This gives the following result.

Corollary 4.4. *Let $D : J \to \mathbf{Sets}^{\mathcal{C}^{op}}$ be a diagram of type J. A cone $\theta_j : F \to D_j$ is the limit (or colimit) of D if and only if $(\theta_j)_C : F(C) \to D_j(C)$ is the limit (or colimit) of* $\mathrm{ev}_C \circ D$ *for each $C \in \mathcal{C}$. Here* ev_C *is the functor mapping $G \in \mathbf{Sets}^{\mathcal{C}^{op}}$ to $G(C)$.*

Proof. The "only if" direction follows directly from 4.3, and the fact that right adjoints and left adjoints preserves limits and colimits respectively.

For the other direction, we simply note that for any other cone $\theta'_j : G \to D_j$, $(\theta'_j)_C$ factors through $(\theta_j)_C$, so θ'_j must factor through θ_j. □

For the interpretation of logic, we are interested in how to construct subobjects of a given presheaf. Since presheaves consist of a collection of sets which behaves well with respect to a collection of maps, it is not hard to believe that subobjects in such a category are given by presheaves which pointwise form subsets.

Proposition 4.5. *Given a presheaf $F : \mathcal{C} \to \mathbf{Sets}$, an arrow $a : G \to F$ is mono if and only if $a_C : G(C) \to F(C)$ is injective for all $C \in \mathcal{C}$. Furthermore, for $a : G \rightarrowtail F, a' : G' \rightarrowtail F$, G and G' are equivalent if and only if $\mathrm{im}(a_C) = \mathrm{im}(a'_C)$ for all $C \in \mathcal{C}$.*

Proof. By Corollary 4.4 and the fact that monos can be characterized by limits, the first statement follows directly.

For the second statement, we note that using Proposition 3.21, we can characterize equivalences of subobjects entirely in terms of monos, epis and commutative diagrams. Thus, by 4.4, two monos are equivalent if and only if they are equivalent pointwise for each C. But then the result is obvious. □

Remark 4.6. The last part of the proposition allows us to, given a subobject $a : G \to F$, construct the functor $F_a : \mathcal{C} \to \mathbf{Sets}$ defined by $F(C) = \mathrm{im}(a_C)$ and with inclusions as maps. This will be used as a canonical representative for the equivalence class. We will use the habit of identifying the equivalence class with this representative.

Proposition 4.7. *For a subobject $S \rightarrowtail A$, the characteristic map of S is given by the following mapping. For each $C \in \mathcal{C}$ and each $a \in A(C)$, χ_C is defined by sending a to the set of arrows f with codomain C such that, for $f : D \to C$, $A(f)(a) \in S(D)$.*

Proof. First, we need to show that this set is actually a sieve. But it follows from the naturality of the inclusions that if $g : D' \to D$ and $A(f)(a) \in S(D)$, then $A(c \circ f)(a) \in S(D')$.

Furthermore, we need to show that it is a natural transformation, i.e. we need to show that

$$\begin{array}{ccc} A(C) & \xrightarrow{\chi_C} & \Omega(C) \\ \downarrow{\scriptstyle A(f)} & & \downarrow{\scriptstyle \Omega(f)} \\ A(D) & \xrightarrow{\chi_D} & \Omega(D). \end{array} \tag{4.2}$$

commutes for all $f : D \to C$. Assume that $g \in \chi_D(A(f)(a))$, then it must hold that $f \circ g \in \chi_C(a)$. But then $g \in \Omega(f)(\chi_C(a))$ showing the desired property.

Next, we need to show that this indeed gives the desired pullback square. By Corollary 4.4, it is enough to show that it is the desired pullback square for each C. Thus we consider the diagram

$$\begin{array}{ccccc} T & & & & \\ & \searrow^{f} & \overset{!}{\longrightarrow} & \searrow & \\ & & S(C) & \longrightarrow & 1(C) \\ & & \downarrow \iota & & \downarrow \top_C \\ & & A(C) & \xrightarrow{\chi_C} & \Omega(C). \end{array} \tag{4.3}$$

Assume there is $t \in T$ such that $f(t) \notin S(C)$. Then $\xi_C(f(t))$ does not contain id, although $\top_C(!t))$ does, showing that the diagram can not commute. Thus, whenever the diagram do commute, f factors through $S(C)$. Uniqueness follows from the fact that ι is monic. ☐

We will now make explicit the Heyting algebra of Theorem 3.19 for the case of presheaf topoi.

Proposition 4.8 ([16] I.8.5)**.** *Given a small category $\mathcal{C}$, and an object $A \in \mathbf{Sets}^{\mathcal{C}^{op}}$, the operations described in Theorem 3.19 can be described explicitly for (canonical representatives of) subobjects $T \rightarrowtail A, S \rightarrowtail A$ in the following way:*

(i) The meet is given by $S \wedge T(C) = S(C) \cap T(C)$

(ii) The join is given by $S \vee T(C) = S(C) \cup T(C)$.

(iii) Implication is given by

$$S \Rightarrow T(C) = \{x \in A(C) \mid \forall D \in \mathcal{C}, \forall f : D \to C, A(f)(x) \in S(D) \Rightarrow A(f)(x) \in T(D)\} \tag{4.4}$$

(iv) Negation is given by

$$\neg S(C) = \{x \in A(C) \mid \forall D \in \mathcal{C}, \forall f : D \to C, A(f)(x) \notin S(D)\} \tag{4.5}$$

Note that the negation is only a rewriting of the definition of negation by $S \Rightarrow \varnothing$.

Proof. Since for each C,

$$\begin{array}{ccc} T(C) \cap S(C) & \longrightarrow & S(C) \\ \downarrow & & \downarrow \\ T(C) & \longrightarrow & A(C) \end{array} \tag{4.6}$$

is a pullback diagram in **Sets**, (i) follows from Proposition 4.4. A similar argument holds for the join.

For implication, suppose that the characteristic map of S is denoted s and the characteristic map of $S \cap T$ is denoted $s \cap t$, then we want to consider the diagram

$$E(C) \rightarrowtail A(C) \underset{(s \cap r)_C}{\overset{s_C}{\rightrightarrows}} \Omega(C). \tag{4.7}$$

Being an equalizer in **Sets** means that $s_C(a) = (s \cap t)_C(a)$ for all $a \in E(C)$. From Proposition 4.7, the action of s_C and $(s \cap t)_C$ is known. Thus, we are looking for elements a with the property that $f : D \to C$ is such that $A(f)(a) \in S(D)$ if and only if $A(f)(a) \in S(D) \cap T(D)$. But this gives exactly the set defined above. Since the negation is just a special case of the implication, the proof is done. □

Proposition 4.9. *Any presheaf topos* $\mathbf{Sets}^{\mathcal{C}^{op}}$ *is geometric with the arbitrary join given for a family* $\{S_i\}$ *of (canonical representatives of) subobjects of* A *by* $\left(\bigvee_i S_i\right)(C) = \bigcup_i S_i(C)$.

Proof. We define $S : \mathcal{C}^{\mathrm{op}} \to \mathbf{Sets}$ by $S(C) = \bigcup_i S_i(C)$ on objects and by the restriction of $A(f)$ on arrows. This is clearly a well defined functor and a subobject of A. It is also clear that each inclusion $S_i \to A$ factors through S. Finally, we note that if each inclusion $S_i \to A$ also factors through S', then $\bigcup_i S_i(C) \subseteq S'(C)$, so the inclusion $S \to A$ factors through S'. □

One way to reason about logic in presheaf categories is through the use of forcing which we will now introduce. Later we will talk about sheaf categories, and we note that the same concept of forcing makes sense in these categories.

Definition 4.10 ([16] §VI.7)**.** Let $\mathcal{C}$ be a small category, and consider the topos $\mathbf{Sets}^{\mathcal{C}^{\mathrm{op}}}$. For an object $C \in \mathcal{C}$ we say that C *forces* $\varphi[\vec{s}/\vec{x}]$, written $C \Vdash \varphi[\vec{s}/\vec{x}]$, if given that the type of $\vec{x}$ is interpreted by X, the map $[\![\vec{s}]\!] : yC \to X$ factors through $[\![\vec{x}.\varphi]\!]$. If we, by Yoneda, interpret $[\![\vec{s}]\!]$ as an element of $X(C)$, this is the same as saying that $[\![\vec{s}]\!] \in [\![\vec{x}.\varphi]\!](C)$.

Let us now consider the relationship between forcing and satisfaction of formulae as defined in 3.26. We note that for a formula in a context interpreted as X to be valid, it needs to be all of X. Thus the following proposition follows immediately.

Proposition 4.11. *Let* $x.\varphi$ *be a formula in the Mitchell–Bénabou language, and let the type of the context be* $X \in \mathcal{T}$. *Then* $\mathcal{T} \vDash \vec{x}.\varphi$ *if and only if for any* C *and any* $[\![\vec{s}]\!] \in X(C)$, $C \Vdash \varphi[\vec{s}/\vec{x}]$.

The following two theorems will be a great help when determining the validity of formulas in a presheaf topos. There is also a sheaf version of this theorem which will be presented in Section 4.2.

We first note that if $C, D \in \mathcal{C}$ for some small category $\mathcal{C}$, then each arrow $f : D \to C$ induces a natural transformation $yf : yD \to yC$ by post-composition. Thus, for each term $\vec{s}$ in the Mitchell–Bénabou language, there is also a term $\vec{s} \circ yf : yD \to X$. To shorten the notation, we will call this term $f^*(\vec{s})$.

Theorem 4.12 ([16] §VI.7)**.** *Let* $C \Vdash \varphi[\vec{s}/\vec{x}]$, *and let* $f : D \to C$. *Then* $D \Vdash \varphi[f^*(\vec{s})/\vec{x}]$. *This property is called monotonicity.* □

Theorem 4.13 ([16] §VI.7)**.** *Let* $\mathcal{E} = \mathbf{Sets}^{\mathcal{C}^{op}}$ *for some small* $\mathcal{C}$, *and let* $\vec{x}.\varphi$ *and* $\vec{x}.\psi$ *be formulas. Then*

(i) $C \Vdash (\varphi \wedge \psi)[\vec{s}/\vec{x}]$ *if and only if* $C \Vdash \varphi[\vec{s}/\vec{x}]$ *and* $C \Vdash \psi[\vec{s}/\vec{x}]$.

(ii) $C \Vdash (\varphi \vee \psi)[\vec{s}/\vec{x}]$ *if and only if* $C \Vdash \varphi[\vec{s}/\vec{x}]$ *or* $C \Vdash \psi[\vec{s}/\vec{x}]$.

(iii) $C \Vdash (\varphi \Rightarrow \psi)[\vec{s}/\vec{x}]$ *if and only if for all* $f : D \to C$, $D \Vdash \varphi[f^*(\vec{s})/\vec{x}]$ *implies* $D \Vdash \psi[f^*(\vec{s})/\vec{x}]$.

(iv) $C \Vdash (\neg\varphi)$ *if and only if there is no* $f : D \to C$ *such that* $D \Vdash \psi[f^*(\vec{s})/\vec{x}]$.

Assume next that $\vec{x}y.\varphi$ *has at least the free variable* $y : Y$. *Then*

(v) $C \Vdash (\exists y\varphi)[\vec{s}/\vec{x}]$ *if and only if there exists a* $t \in Y(C)$ *such that* $C \Vdash \varphi[\vec{s}/\vec{x}, t/y]$.

(vi) $C \Vdash (\forall y\varphi)[\vec{s}/\vec{x}]$ *if and only if for all* $f : D \to C$ *and all* $t \in Y(D)$ *we have* $D \Vdash \varphi[f^*(\vec{s})/\vec{x}, t/y]$. ☐

We know that the law of excluded middle and the axiom of choice does not in general hold in a presheaf topos. However, we will later need a slight weakening of the axiom of choice, called the axiom of dependent choice. This axiom does, in fact, hold in any presheaf topos. To show this, we start with a lemma.

Lemma 4.14. *Let* $\mathcal{T}$ *be a topos with a natural numbers object* $\mathbb{N}$. *Then for* $A \in \mathcal{T}$, $\mathbb{N} \times A \cong A + A + \cdots$, *where the latter is a countable coproduct of the objects* A.

Proof. Denote by ι_i the inclusion of A into the ith term in the coproduct. By the universal property of a coproduct, it is possible to construct a map, call it S making the diagram

$$\begin{array}{ccc} A & \xrightarrow{\iota_i} & A + A + \cdots \\ & \searrow^{\iota_{i+1}} & \downarrow S \\ & & A + A + \cdots . \end{array} \tag{4.8}$$

commute. We note that $\iota_n = \underbrace{S \circ \cdots \circ S}_{n \text{ times}} \circ \iota_0$. By the universal property of the natural numbers object, there exists a unique h such that

$$\begin{array}{ccccc} A & \xrightarrow{(0,\mathrm{id})} & N \times A & \xrightarrow{(s,\mathrm{id})} & N \times A \\ & \searrow^{\iota_1} & \downarrow h & & \downarrow h \\ & & A + A + \cdots & \xrightarrow{S} & A + A + \cdots . \end{array} \tag{4.9}$$

But this means that each $S^n \circ \iota_0 = \iota_n$ factors through $N \times A$. Next, we note that we have a family of maps $(s^n, \mathrm{id}) \circ (0, \mathrm{id}) : A \to N \times A$. Thus, we have the commutative diagram

$$\begin{array}{ccccc} A & \xrightarrow{\iota_1} & A + A + \cdots & \xrightarrow{S} & A + A + \cdots \\ & \searrow^{(0,\mathrm{id})} & \downarrow g & & \downarrow g \\ & & N \times A & \xrightarrow{(s,\mathrm{id})} & N \times A. \end{array} \tag{4.10}$$

But from the commutativity of these diagrams, it follows that $\iota_n = \iota_1 \circ S^n \circ h \circ g = \iota_n \circ h \circ g$. But then it follows from the coproduct property that $h \circ g = \mathrm{id}$. ☐

Theorem 4.15 (cf. [16] VI Exercise 12)**.** *The axiom of dependent choice, i.e. the formula*

$$\forall x \exists y R(x, y) \Rightarrow \forall x \exists f \in X^{\mathbb{N}}(f(0) = x \wedge \forall n R(f(n), f(n+1)) \tag{4.11}$$

holds in all presheaf topoi for all objects X *and relations* $R \to X \times X$.

Proof. We note that the formula is over the empty context, which is always interpreted by 1. By Proposition 4.11 we want to show that C' forces the formula for each C' in the base category.

By Theorem 4.13 we want to show

$$C \Vdash \forall x \exists f \in X^{\mathbb{N}}(f(0) = x \wedge \forall n R(f(n), f(n+1)) \tag{4.12}$$

assuming $C \Vdash \forall x \exists y R(x, y)$ for all $C \to C'$. Since C' was arbitrary, this last condition does however not say anything of interest.

The assumption $C \Vdash \forall x \exists y R(x, y)$ can be rewritten as saying for all $D \to C$ and all $a \in X(D)$, there exists $b \in X(D)$ such that $(a, b) \in R(D)$. Fixing D, this is thus exactly the hypothesis for the regular axiom of dependent choice.

By (external) dependent choice, we thus have that for all $\alpha \in X(D)$, there exists a $g \in X(D)^{\mathbb{N}}$ such that $f(0) = \alpha$ and for all $n \in \mathbb{N}$, $(g(n), g(n+1)) \in R(D)$. Suppose now that we have $a : D' \to D$. Then $a^* \circ g \in X(D')^{\mathbb{N}}$.

Using Lemma 4.14, we can write

$$\begin{aligned} X(D')^{\mathbb{N}} = \operatorname{Hom}(yD', X)^{\mathbb{N}} = \operatorname{Hom}(yD', X) \times \operatorname{Hom}(yD', X) \times \cdots = \\ = \operatorname{Hom}(yD' + yD' + \cdots, X) = \operatorname{Hom}(yD' \times \mathbb{N}, X) = X^{\mathbb{N}}(D'). \end{aligned} \quad (4.13)$$

Thus we have enough to conclude that $D \Vdash (\forall n \in \mathbb{N}(R(f(n), f(n+1))))[g/f]$ where f is a variable with its type interpreted by $X^{\mathbb{N}}$. But now, the rest of the proof is just a straightforward application of the "only if"-direction of Theorem 4.13. □

4.2 Sheaf Topoi

Before introducing sheaves, we will give a brief introduction to the concept of a locale. Locales are related to spaces, or perhaps even more to the concept of the opens of a space. We know that any continuous map $f : X \to Y$ induces a map $f^* : \mathcal{O}(Y) \to \mathcal{O}(X)$. For locales, we start in the other direction, defining maps between generalisations of sets of opens.

Definition 4.16 ([29] C1.1.1)**.** A *frame* is a complete lattice, that is a lattice with arbitrary meets and joins, for which the distributive law

$$a \wedge \bigvee_{i \in I} b_i = \bigvee_{i \in I} (a \wedge b_i) \quad (4.14)$$

holds for arbitrary index sets I.

A *homomorphism of frames* is a map preserving finite meets and arbitrary joins.

Definition 4.17 ([29] C1.2.1)**.** A *locale* is a frame. A *homomorphism of locales* is an arrow in the opposite category of the category of frames.

Remark 4.18. Note that any topology gives rise to a frame by taking the set of its opens ordered by inclusion. We define the functor $\mathcal{O} : \mathbf{Top} \to \mathbf{Loc}$ by sending each topological space to the frame of its locales, and each continuous map to the induced frame map of opens.

We will now turn to the introduction of sheaves, which will be a very important concept for the coming sections. First, we need the definition of an open cover.

Definition 4.19. Given a locale L, an *open covering* of an element $U \in L$ is a set $C = \{U_i \in F \mid U_i \leq U\}$ such that $\bigvee C = U$.

Definition 4.20 ([16] §II.1)**.** A *sheaf* is a presheaf $F : L \to \mathbf{Sets}$ where L is a locale (for instance the set of opens of a topological space) such that for each open set $U \in L$, and for each covering $\{U_i\}$ of U, the following digram

$$FU \xrightarrow{e} \prod_i FU_i \underset{q}{\overset{p}{\rightrightarrows}} \prod_{i,j} F(U_i \cap U_j) \quad (4.15)$$

is an equalizer (in **Sets**). Here $e(t) = \{t|_{U_i}\}, p(\{t_i\}) = \{t_i|_{(U_i \cap U_j)}\}$ and $q(\{t_j\}) = \{t_j|_{(U_i \cap U_j)}\}$. The category of sheaves on L is called $\mathbf{Sh}(L)$. If L is the opens of a space X, we usually write $\mathbf{Sh}(X)$. Note that this is a full subcategory of the category of presheaves on L.

Let us consider the meaning of the above definition. An equalizer in **Sets** is the set of all elements x for which $p(x) = q(x)$. Thus, this says that any set $\{t_i\}$ of functions on the sets U_i which are equal on all intersections comes from a unique function on all of U. This is often phrased as the locality and glueing axioms, which are often used to define sheaves. To describe these, we let L be a locale, $U \in L$ an open, and (U_i) an open cover of U. We furthermore assume that F is a presheaf on L. A sheaf can then be characterized by the following axioms:

- (Locality) If $f, g \in F(U)$ and $f|_{U_i} = g|_{U_i}$ for each U_i, then $g = f$.
- (Gluing) If there is a collection $f_i \in F(U_i)$ with the property that $f_i|_{U_i \cap U_j} = f_j|_{U_i \cap U_j}$ for all U_i, U_j, then there is an $f \in F(U)$ such that $f|_{U_i} = f_i$ for each U_i.

We note that the gluing property is the existence mentioned above, and that the locality is the uniqueness.

In the later parts of the thesis, we will mainly be interested in categories where the underlying locale comes from a topological space. However, we note that in most of the following, there are corresponding results for locales.

Proposition 4.21 ([16] II.2.2)**.** *The category of sheaves on a space X has all small limits. Furthermore, the limit is exactly the limit of these sheaves seen as presheaves.*

Theorem 4.22 ([16] II.5.4)**.** *Let $\iota : \mathbf{Sh}(X) \to \mathbf{Sets}^{\mathcal{O}(X)^{op}}$ be the inclusion functor. Then ι has a left adjoint a called the* associated sheaf functor. ☐

For any presheaf A which is already a sheaf, the associated sheaf functor works as the identity. Since this functor is a left adjoint, it preserves colimits. This means that any colimit in $\mathbf{Sh}(X)$ can be calculated by considering the colimit of the diagram of sheaves as a diagram of presheaves, and then apply a to the resulting colimit. In particular, this means that $\mathbf{Sh}(X)$ has all small colimits.

Proposition 4.23 ([16] II.8.1)**.** *The category $\mathbf{Sh}(X)$ has exponentials, where the exponential of two sheaves is exactly the presheaf exponential of the same. I.e.*

$$F^G(U) = \hom(yU \times G, F). \tag{4.16}$$

☐

Proposition 4.24 (II.8.2)**.** *The functor Ω on $\mathcal{O}(X)$, defined on objects by*

$$\Omega U = \{V | V \in \mathcal{O}(X), V \subseteq U\} \tag{4.17}$$

and on inclusions $W \subseteq U$ by mapping $V \mapsto V \cap W$, is a sheaf on X, and a subobject classifier for the category $\mathbf{Sh}(X)$. ☐

Summing these last propositions up gives the following result.

Theorem 4.25 ([16] §II.8)**.** *The category $\mathbf{Sh}(X)$ is a topos.*

We also have the following result.

Proposition 4.26 ([29] §1.4)**.** *The category $\mathbf{Sh}(X)$ of sheaves over a space X is geometric.*

There is a sheaf version of Theorem 4.13.

Theorem 4.27 ([16] §VI.7)**.** *A modified version of Theorem 4.13 holds for sheaf topoi. The modification goes as follows. Let $\vec{x}.\varphi$ and $\vec{x}.\psi$ be formulas. Then*

(i) $U \Vdash (\varphi \wedge \psi)[\vec{s}/\vec{x}]$ *if and only if* $U \Vdash \varphi[\vec{s}/\vec{x}]$ *and* $U \Vdash \psi[\vec{s}/\vec{x}]$.

(ii) $U \Vdash (\varphi \vee \psi)[\vec{s}/\vec{x}]$ *if and only if there is a cover* $\{U_i\}$ *of* U *such that* $U_i \Vdash \varphi[\vec{s}|_{U_i}/\vec{x}]$ *or* $U_i \Vdash \psi[\vec{s}|_{U_i}/\vec{x}]$ *for all* U_i.

(iii) $U \Vdash (\varphi \Rightarrow \psi)[\vec{s}/\vec{x}]$ *if and only if for all* $V \to U$, $V \Vdash \varphi[\vec{s}|_V/\vec{x}]$ *implies* $V \Vdash \psi[\vec{s}|_V/\vec{x}]$.

(iv) $U \Vdash (\neg\varphi)$ *if and only if only* $V = \varnothing$ *forces* $\varphi[\vec{s}|_V/\vec{x}]$.

Assume next that $\vec{x}y.\varphi$ *has at least the free variable* $y : Y$*. Then*

(v) $U \Vdash (\exists y\varphi)[\vec{s}/\vec{x}]$ *if and only if there exists a covering* $\{U_i\}$ *and* $t_i \in Y(U_i)$ *such that* $U_i \Vdash (\varphi)[\vec{s}|_{U_i}/\vec{x}, t_i/y]$.

(vi) $U \Vdash (\forall y\varphi)[\vec{s}/\vec{x}]$ *if and only if for all* $V \leq U$ *and all* $t \in Y(V)$ *we have* $V \Vdash (\varphi)[\vec{s}|_V/\vec{x}, t/y]$. □

Remark 4.28. The monotonicity (Theorem 4.12) holds also for presheaves. Furthemore, there is an important special case of (ii) above. Namely, let $\psi = \perp$. Then the rule reads $U \Vdash \varphi[\vec{s}/\vec{x}]$ if and only if there is a cover $\{U_i\}$ of U such that $U_i \Vdash \varphi[\vec{s}|_{U_i}/\vec{x}]$ for all U_i. This property is called *local character.*

The full set of opens of a topological space is usually a rather complicated construction, and it is in many cases helpful to just look at a basis of the space. Fortunately, sheaves behave well with respect to this.

Proposition 4.29 ([16] II.1.3)**.** *Given a topology $\mathcal{T}$ on X and a basis $\mathcal{B}$ for this topology, the restriction functor gives an equivalence of categories* $\mathbf{Sh}(X) \cong \mathbf{Sh}(\mathcal{B})$*. In other words, it is possible to define a sheaf by giving a definition only on the basis of the topology.*

4.3 Geometric Morphisms

As is customary when having introduced new mathematical objects, we now want to introduce arrows between them. Here, we will introduce geometric morphisms, one type of morphism between topoi. We have delayed our introduction of this morphism until after introducing categories of sheaves since there is a very tight connection between geometric morphisms of such topoi and the morphisms of their base categories.

Definition 4.30 ([16] VII.1)**.** Let $\mathcal{E}, \mathcal{F}$ be topoi. A pair of functors $g^* : \mathcal{E} \to \mathcal{F}, g_* : \mathcal{F} \to \mathcal{E}$, is called a *geometric morphism* g if g^* is left adjoint to g_* and g^* preserves small limits. This is often written as $g : \mathcal{F} \to \mathcal{E}$. The map g_* is called the *direct image* of the geometric morphism, and the functor g^* is called the inverse image.

An interesting property of geometric morphisms is that their inverse images preserve geometric formulae. In order to describe this, we will first note that given a geometric morphism (or in fact any functor preserving small limits) there is a canonical way to compare interpretations.

Definition 4.31 (cf. [16] §X.2)**.** Let M be a Σ-structure in $\mathcal{F}$. Given a functor $f : \mathcal{F} \to \mathcal{E}$ preserving limits, we let the *induced Σ-structire* in $\mathcal{E}$ be the structure given by the following assignments (cf. Definition 3.22):

1. For each sort A, the object $f(MA)$.
2. For a list of sorts $A_1, \cdots, A_n$, the object $f(MA_1 \times \cdots \times MA_n) = f(MA_1) \times \cdots \times f(MA_n)$.
3. For each function symbol $g : A_1 \cdots A_n \to B$, the arrow $f(Mg)$.
4. For each relation $R \rightarrowtail A_1 \cdots A_n$, the subobject $f(MR) \rightarrowtail f(MA_1 \times \cdots \times A_n)$.

Theorem 4.32 ([16] X.3.1)**.** *Let $f : \mathcal{F} \to \mathcal{E}$ be a geometric morphism and let φ be a geometric formula in some context. Then $[\![\varphi]\!]_{f^*(M)} = f^*([\![\varphi]\!]_M)$.* □

What this theorem says is that interpretations are preserved under geometric morphisms. In particular, it means that if a sequent or formula is satisfied in M, then it will also be satisfied in $f^*(M)$.

Given sheaf categories, morphisms of the underlying "geometric" categories induce geometric morphisms. We want to consider these morphisms in some more detail, for which we need the introduction of bundles and étalé spaces.

Definition 4.33 ([16] §II.4)**.** A *bundle over X* is a continuous map $E \to X$ of spaces. A *morphism of bundles* is a commutative diagram

$$\begin{array}{ccc} E & \longrightarrow & E' \\ & \searrow & \downarrow \\ & & X. \end{array} \tag{4.18}$$

In other words, the category of bundles over X is exactly the slice category $\mathbf{Top}/X$ where $\mathbf{Top}$ is the category of topological spaces.

Definition 4.34 ([16] II.6)**.** A continuous map $f : E \to X$ between spaces is a *local homeomorphism* if for each $e \in E$ there is an open set U containing e such that $f(U)$ is open in X and that the restriction of f to U is a homeomorphism between U and $f(U)$. A bundle for which the map is a local homoemorphism is called an *étalé space*. The collection of étalé spaces over X together with bundle morphisms gives the category $\textbf{étalé}(X)$.

A sheaf is a space where each open is sent to some set. An étalé space is a space mapping continuously to an underlying set. Thus, both structures associate something to each open of a set, in light of which the following might not be too surprising.

Proposition 4.35 ([16] II.6.3)**.** *For a space X, the categories $\mathbf{Sh}(X)$ and $\textbf{étalé}(X)$ are equivalent. The functor $\Gamma : \textbf{étalé}(X) \to \mathbf{Sh}(X)$ is given on $f : F \to X$ by associating each U to the set of commutative diagrams*

$$\begin{array}{ccc} U & \longrightarrow & F \\ & \searrow & \downarrow{\scriptstyle f} \\ & & X. \end{array} \tag{4.19}$$

□

Thus we can consider our objects as sheaves or étalé spaces, depending on which approach works best for the situation. This gives us a way to describe the geometric morphism promised above.

Theorem 4.36 ([16] II.9.1 II.9.2)**.** *A continuous map $f : Y \to X$ induces both a functor $f_* : \mathbf{Sh}(Y) \to \mathbf{Sh}(X)$ and $f^* : \mathbf{Sh}(X) \to \mathbf{Sh}(Y)$. The first one is given by precomposition, i.e.*

$f_*F(U) = F(f^{-1}U)$. *The second one is given by seeing a sheaf* F *as an étalé space* $g : E \to X$ *and taking the pullback*

$$\begin{array}{ccc} E' & \longrightarrow & E \\ \downarrow & & \downarrow g \\ Y & \xrightarrow{f} & X. \end{array} \tag{4.20}$$

Furthermore, these maps constitutes the direct image and inverse image of a geometric morphism. □

Remark 4.37. Note that it is not a priori clear that the presheaves and the bundles defined by f_* and f^* are sheaves and étalé spaces respectively. This is part of the content of the theorem.

4.4 Atomic Geometric Morphisms

In this section, we would like to introduce the concept of an atomic geometric morphism. Such a morphism is one where the logical structure of the topos is preserved. We will also discuss some other results that will help us find atomic geometric morphisms.

Definition 4.38 ([16] §IV.2)**.** Let $F : \mathcal{F} \to \mathcal{E}$ be a functor between topoi. Then F is a *logical morphism* if it preserves the structure used in defining a topos, that is finite limits, exponentials and subobject classifiers.

Definition 4.39 ([29] §C3.5)**.** Let $g : \mathcal{F} \to \mathcal{E}$ be a geometric morphism, and assume that g^* is a logical morphism. Then g is called *atomic*.

We will now introduce two concepts that can be used to find atomic geometric morphisms. For the first concept we mention the result that given a topos $\mathcal{E}$, the slice category $\mathcal{E}/F$ for $F \in \mathcal{E}$ is also a topos ([16] IV.7.1).

Definition 4.40 ([29] §C3.3)**.** Let $\mathcal{E}$ be a topos, and let $F \in \mathcal{E}$. Let $\pi : \mathcal{E}/F \to \mathcal{E}$ be the geometric morphism defined by π^* being the pullback along the unique map $F \to 1$, and π_* being its right adjoint. If a geometric morphism $\mathcal{F} \to \mathcal{E}$ factors as $\mathcal{F} \xrightarrow{\cong} \mathcal{E}/F \xrightarrow{\pi} \mathcal{E}$, then it is called a *local homeomorphism of topoi.*

Definition 4.41 ([29] A4.6.1)**.** Given a geometric morphism $f : \mathcal{E} \to \mathcal{F}$, the morphism is *localic* if every $E \in \mathcal{E}$ is a subobject of some coequalizer $A \to f^*(F)$ for some $F \in \mathcal{F}$.

Proposition 4.42 ([29] C3.5.4)**.** *A localic morphism is atomic if and only if it is a local homeomorphism.*

Since we are working primarily with sheaves, we will now try to relate these concepts to sheaves.

Lemma 4.43 ([29] §A4.6)**.** *If* $f : X \to Y$ *is a continous map, then the induced geometric morphism is localic.* □

Lemma 4.44. *Any local homeomorphism of spaces* $f : Y \to X$ *induces an atomic geometric morphism.*

Proof. By Proposition 4.42 and Lemma 4.43 it is enough to show that the induced geometric morphism is a local homeomorphism.

Since $Y \to X$ is a local homeomorphism, it represents some sheaf in $\mathbf{Sh}(X)$, say F. We will now show that $f : \mathbf{Sh}(Y) \to \mathbf{Sh}(X)$ factors as $\mathbf{Sh}(Y) \xrightarrow{\cong} \mathbf{Sh}(X)/F \xrightarrow{\pi} \mathbf{Sh}(X)$. We start by noting that any object in $\mathbf{Sh}(X)/F$ can be represented by a commutative diagram

$$\begin{array}{ccc} A & \xrightarrow{a} & Y \\ & \searrow^{a'} & \downarrow f \\ & & X \end{array} \tag{4.21}$$

where a and a' are local homeomorphisms. Clearly, there is a bijective correspondence between such diagrams and local homeomorphisms into Y. We thus have $\mathbf{Sh}(Y) \cong \mathbf{Sh}(X)/F$. We see that $\mathrm{id} : X \to X$ is the $\textbf{étalé}(X)$ representation of the terminal object in $\mathbf{Sh}(X)$, But pullback along the diagram

$$\begin{array}{ccc} Y & \xrightarrow{f} & X \\ & \searrow^{f} & \| \\ & & X \end{array} \tag{4.22}$$

is just pullback along $f : Y \to X$, showing that f indeed factors as required. $\square$

4.5 Local truth

Definition 4.45. Given a space X, a context $\vec{x}$, a formula φ and an open subset $U \subseteq X$, we will say that $\vec{x}.\varphi$ is *valid at U* if $\iota^*([\![\varphi]\!]) = \iota^*([\![\vec{x}.\top]\!])$. If, for each $x \in X$ there is an open V such that φ is valid at V, we say that φ is *locally valid*.

This description of local validity doesn't really provide anything new. We will show shortly that it is exactly forcing in a new guise. For this, we will need the following lemma.

Lemma 4.46. *Any representable presheaf is a sheaf. Furthermore, the étalé space associated with yU for an open $U \subseteq X$, is exactly the inclusion $\iota : U \rightarrowtail X$.*

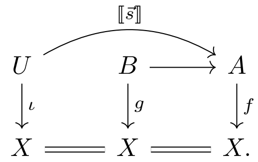
Proof. The locality and glueing axioms are obvious. It is also clear that $\iota : U \to X$ has the property that there exists a unique commutative diagram

$$\begin{array}{ccc} V & \longrightarrow & U \\ & \searrow^{\iota} & \downarrow \iota \\ & & X \end{array} \tag{4.23}$$

exactly when $V \leq U$. $\square$

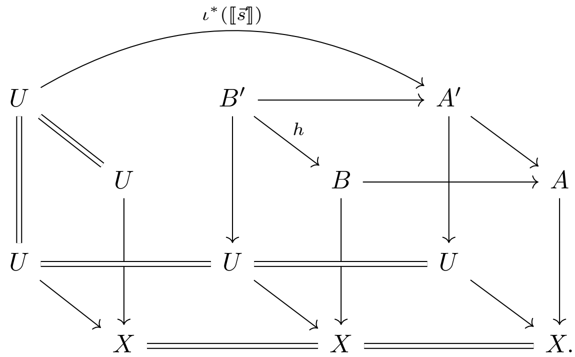
Theorem 4.47. *Let $\mathbf{Sh}(X)$ be a sheaf topos on a space X, $U \subseteq X$, and $\vec{x}.\varphi$ a first order or geometric formula where $\vec{x}$ has types interpreted by $\tilde{A}$, then $U \Vdash \varphi[\vec{s}/\vec{x}]$ in $\mathbf{Sh}(X)$ if and only if $1 \Vdash \varphi[\iota(\vec{s})/\vec{x}]$ in $\mathbf{Sh}(U)$. In particular any closed formula φ is valid at U if and only if U forces φ.*

Proof. Let $f : A \to X$ and $g : B \to X$ be the étalé spaces corresponding to the sheaves $\tilde{A}$ and $[\![\varphi]\!]$ respectively. By Lemma 4.46, we then want consider the diagram



(4.24)

and the question of existens of an arrow $U \to B$ making it commute. Applying ι^*, and remembering that pullback is functorial, we get the diagram



(4.25)

It is immediately clear that if there exists a factorization making the front part commute, then so does the back (by functorality). Since formulas are preserved by ι^*, this means that if $U \Vdash \varphi[\vec{s}/\vec{x}]$, then $1 \Vdash \varphi[\iota^*(\vec{s})/\vec{x}]$ in $\mathbf{Sh}(U)$. On the other hand, suppose that there is an arrow $\beta : U \to B'$ making the back rectangle commute. Then $\beta \circ h$, where h is as defined in the diagram above, is an arrow between U and B, making the front rectangle commute. Thus there exists a map making $[\![\vec{s}]\!]$ factor through through B. ☐

Corollary 4.48. *A locally valid formula is valid.*

Proof. This follows directly from the local character property, introduced in Remark 4.28 ☐

We will now introduce a lemma that will help us to prove some more results.

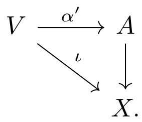
Lemma 4.49. *Let X be a space, and let $V \leq U$. If*

$$\begin{array}{ccc} A' & \xrightarrow{\gamma} & A \\ \downarrow & & \downarrow \\ U & \xrightarrow{\iota} & X \end{array} \tag{4.26}$$

is a pullback diagram, then there is a bijective correspondence between diagrams on the form

$$\begin{array}{ccc} V & \xrightarrow{\alpha} & A' \\ & \searrow^{\iota} & \downarrow \\ & & U \end{array} \tag{4.27}$$

and diagrams

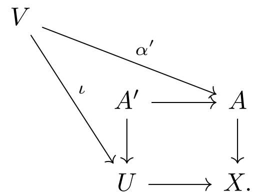


(4.28)

Furthermore, any such α' factors through α.

Proof. Clearly, any diagram on the first form gives rise to a diagram of the second, by just composing with the pullback square in 4.26.

Now, given a diagram of the second form, we note that the inclusion of V into X clearly factors through U. Thus, we have the following diagram

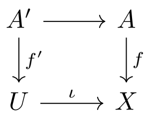


(4.29)

By the pullback property of the diagram 4.26 it follows that there exists a unique $\alpha : V \to A'$ making the diagram commute. Thus we have obtained a diagram on the first form, and shown that α' factors through α. □

Proposition 4.50. *Given a geometric or first order formula φ, if $V \leq U$, then the sets $[\![\vec{x}.\varphi]\!]_{\mathbf{Sh}(X)}(V)$ and $[\![\vec{x}.\varphi]\!]_{\mathbf{Sh}(U)}(V)$ are isomorphic as sets.*

Proof. Let $f : A \to X$ be the representative of $[\![\vec{x}.\varphi]\!]_{\mathbf{Sh}(X)}$ as an étalé space. We know that for $\iota : U \to X$, the induced geometric morphism preserves interpretations, so $[\![\vec{x}.\varphi]\!]_{\mathbf{Sh}(U)}$ is interpreted by some $f' : A' \to U$ making the square



(4.30)

a pullback. Since $[\![\vec{x}.\varphi]\!]_{\mathbf{Sh}(X)}(V) = \mathrm{Hom}_{\mathbf{étalé}(X)}(V, A)$ and $[\![\vec{x}.\varphi]\!]_{\mathbf{Sh}(U)}(V) = \mathrm{Hom}_{\mathbf{étalé}(U)}(V, A')$, the result follows directly from Lemma 4.49. □

Remark 4.51. In the case when $\varphi = \vec{x}.\top$ where the type of $\vec{x}$ is interpreted by X, the Proposition says that $\iota(X)(V) = X(V)$.

4.6 Internal locales

Many of the construction which we will use later will be locales internal to some topos, so we will dedicate this section to such structures. We start by showing that it is possible to give external descriptions of such locales.

Definition 4.52 ([29] §C1.6)**.** Let $f : Y \to X$ be an arrow between locales, and denote by f^{-1} the corresponding frame map. Define the *internalization of* f to be the presheaf $\mathcal{I}(f) = f_*(\Omega_{\mathbf{Sh}(Y)})$, i.e. be defined by

$$\mathcal{I}(f)(U) = \Omega_{\mathbf{Sh}(Y)}(f^{-1}(U)). \tag{4.31}$$

Lemma 4.53 ([29] §C1.6)**.** *Let $f : Y \to X, g : Z \to X, h : Y \to Z$ be arrows between locales such that*

$$\begin{array}{ccc} Y & \xrightarrow{h} & Z \\ & \searrow^{f} & \downarrow g \\ & & X \end{array} \tag{4.32}$$

commutes. Then $\mathcal{I}(h) : \mathcal{I}(g) \to \mathcal{I}(f)$ defined on $V \in \Omega_{\mathbf{Sh}(Y)}(f^{-1}(U))$ by

$$\mathcal{I}(h)_U(V) = h^{-1}(V) \tag{4.33}$$

is a frame map. Thus it gives rise to an internal arrow of locales making $\mathcal{I}$ is a functor $\mathbf{Loc}/X \to \mathbf{Loc}(\mathbf{Sh}(X))$, where $\mathbf{Loc}$ is the category of locales, and $\mathbf{Loc}(\mathbf{Sh}(X))$ denotes the internal category of locales in $\mathbf{Sh}(X)$. □

Theorem 4.54 ([29] C1.6.3)**.** *Then there is an equivalence of categorties*

$$\mathbf{Loc}/X \cong \mathbf{Loc}(\mathbf{Sh}(X)), \tag{4.34}$$

where the functor $\mathbf{Loc}/X \to \mathbf{Loc}(\mathbf{Sh}(X))$ is given by $\mathcal{I}$. □

Remark 4.55. Note that the locale corresponding to $\mathrm{id} : X \to X$ corresponds to $\Omega_{\mathbf{Sh}(X)}$. In the equivalence $\mathbf{Sh}(X) \cong \textbf{étalé}(X)$, $\mathrm{id} : X \to X$ corresponded to 1. Thus, we need to remember that these equivalences behave differently and that it is important to remember which one we are talking about.

Definition 4.56 ([34] §II.1.3)**.** A *point* of a locale X is a frame homeomorphism $p : X \to \Omega$ where Ω is the truth value object. For each $U \in X$, we get a set of points V_U with the property $p(U) = \top$. The collection of such V_U gives a topology on the set of points of X. Thus, there is a functor $\mathrm{pt} : \mathbf{Loc} \to \mathbf{Top}$ mapping X to the topological space of its points.

We will want to describe the internal object of morphisms between two locales. This requires the following lemma.

Lemma 4.57. *Let A be an object. The statement "A is a frame" is a first order statement using the types Ω^A and A. Given two frames A, B, and $f \in B^A$, the statement "f is a frame homomorphism" is first order using the types Ω^A and B^A.*

Proof. The theory of a lattice is clearly first order. The arbitrary meet can be defined as an operation $\bigwedge : \Omega^A \to A$ with the following axioms:

1. $\forall S \in \Omega^A \forall x \in A((x \in S) \Rightarrow (\bigwedge(S) \leq x)$,
2. $\forall S \in \Omega^A \forall x \in A((x > \bigwedge(S)) \Rightarrow (x \in S)$.

and the arbitrary join similarly. The infinite distributive law can be written as $\forall S \forall a((a \wedge \bigwedge(S)) = (\bigwedge(S \cup \{a\}))$ where $S \cup \{a\}$ is the set characterised by the formula $(b \in S) \vee (b = a)$.

For the frame homomorphisms, we first note that given $f \in B^A, S \in \Omega^A$, $f(S)$ can be characterized as $\exists a((a \in S) \wedge (b = f(a)))$. Thus, being a locale homomorphism is characterized by the formulas $f(a \wedge b) = f(a) \wedge f(b)$ and $\bigwedge(f(S)) = f(\bigwedge(S))$. □

Remark 4.58. The above Lemma shows that any theory preserving first-order logic and power objects, thus, any logical morphism and any atomic geometric morphism, preserves the theory of frames.

Theorem 4.59. *Let $\mathcal{I}(f)$ and $\mathcal{I}(g)$ be a locales internal to $\mathbf{Sh}(X)$. Let $\mathcal{L}oc(\mathcal{I}(f), \mathcal{I}(g))$ be the object of internal locale maps. Let $U \subset B$ with $\iota : U \to B$ the inclusion. Define $f_U : A|_U \to U, g_u : B|_U \to U$ to be elements of $\mathbf{Loc}/U$ corresponding to $\iota^*(\mathcal{I}(A)), \iota^*(\mathcal{I}(B))$ respectively. Then the object $\mathcal{L}oc(\mathcal{I}(A), \mathcal{I}(B))$ is given by mapping U to the set of commuting locale diagrams*

$$\begin{array}{ccc} A|_U & \longrightarrow & B|_U \\ & \searrow^{f_U} & \downarrow g_U \\ & & U \end{array} \tag{4.35}$$

and where inclusion $V \subseteq U$ is mapped to the restrictio of the diagram.

Proof. Let α be a variable of type $\mathcal{I}(A)^{\mathcal{I}(B)}$, and let $\alpha.\theta$ be the formula for being a locale homomorphism. By 4.57 we know that this formula is preserved by the geometric morphism ι. What we want to show is that there is a bijective correspondence between $\alpha \in [\![f.\theta]\!]_{\mathbf{Sh}(X)}(U)$ and commutative diagrams of the above form.

From Corollary 4.50 we know that we can just aswell consider the set $[\![f.\theta]\!]_{\mathbf{Sh}(U)}$, where f is now a variable of type $\iota(X^Y) = \iota(X)^{\iota(Y)}$. But $[\![f.\theta]\!]_{\mathbf{Sh}(U)}(U) = \mathrm{Hom}(1, [\![f.\theta]\!]_{\mathbf{Sh}(U)}$, so we are looking for the global elements of $\mathcal{L}oc(\iota^*(X), \iota^*(Y))$ which by the equivalence of categories are exactly given by commutative diagrams of the above form.

By Lemma 4.53, we know that the global elements of $\mathcal{L}oc(\iota^*(X), \iota^*(Y))$ corresponds to frame maps acting like h^{-1}. Now, the restriction maps of presheaf X^Y are just function restrictions. Thus, if we have $\iota : V \subseteq U$, $X^Y(\iota)(h^{-1}) = h^{-1}|_V$, which corresponds exactly to the restrictions of diagrams. □

Corollary 4.60. *The point object of a locale X is given by its local sections, i.e. for each U it is mapped to the set of commutative locale diagrams*

$$\begin{array}{ccc} U & \longrightarrow & L_{X|_U} \\ & \diagdown\!\!\diagdown & \downarrow y_U \\ & & U \end{array} \tag{4.36}$$

Furthermore, the locale has a global point if and only if the point object has a global element.

Proof. Noting that $\mathcal{I}(\iota) = \Omega|_U$, and that $\Omega|_U$ is represented by $\mathrm{id} : U \to U$, the result follows directly from Theorem 4.59. □

Theorem 4.61 ([34] §II.1.4)**.** *There is an adjunction*

$$\hom(\mathcal{O}(X), Y) \cong \hom(X, \mathrm{pt}(Y)). \tag{4.37}$$

Definition 4.62 ([34] §II.1.5)**.** A frame L is called *spatial* if for all $a, b \in L$, if $a \not\leq b$, then there is a point p such that $p(a) = 1$ and $p(b) = 0$.

An important collection of locales are those who are compact completely regular. These will be used in the studies of quantum mechanics later, and we now give a quick introduction to what it means to be compact and completely regular.

Definition 4.63 ([7] 3.1)**.** Let U, V be elements of a frame. We say that U is rather below V, written $U \prec V$, if there exists W such that $U \wedge W = 0$ and $V \vee W = 1$.

Definition 4.64. We say that a frame F is regular if for each $V \in F$, $V = \bigvee \{U \in F | U \prec V\}$.

Definition 4.65 ([35] §0)**.** Let V, U be elements of a frame F, and let $V \ll U$ mean that there is a sequence $(V_{n,k})_{n=1,0,\ldots,;k=0,1,\ldots,2^n}$ such that

$$V = V_{n,0} \qquad V_{n,k} \prec V_{n,k+1} \qquad V_{n,2^n} = U \qquad V_{n,k} = V_{n+1,2k} \tag{4.38}$$

for all n, k. A the frame is called is called *completely regular* if for any $U \in F$, $U = \bigvee\{V \in L \mid V \ll U\}$.

The following definition is standard for topological spaces, but we note that it also applies for locales.

Definition 4.66. A locale X is called *compact* if every open cover, i.e. every collection U_i such that $\bigvee_i U_i = X$, has a finite subcover $U_1, \ldots, U_n$ such that $\bigvee_{j=1}^n U_j = X$.

Although frames and locales usually behave rather well constructively, their construction requires the power object which does not, in general, behave well with geometric morphisms. However, in some cases, it is possible to avoid this problem by defining frames using so-called presentations.

A presentation is given by an object G of generators, and an object R of relations. The informal idea is then to define a collection of formal expressions; the closure over G with respect to formal finite conjunction and formal arbitrary disjunction. Then we factor this collection by provable equality using the relations.

However, a collection containing formal arbitrary disjunctions is generally too large to fit in the topos of G and R.[1] However, there is a way to avoid this, and make some frames behave well with respect to geometric morphisms.

Definition 4.67. The *finite power object* of A, denoted $\mathcal{F}(A)$ is the object of (Kuratowski) finite subobject of A. That is, it is the subobject of Ω^{Ω^A} characterised by the fact that for each B in $\mathcal{F}(A)$, there is a surjective mapping from $\underbrace{1 + \cdots + 1}_{n \text{ times}}$ to B for some n.

Lemma 4.68 ([37] §2)**.** *The finite power object is characterized by a geometric theory.* □

The idea now is to write each relation on the form

$$\bigwedge a \leq \bigvee \bigwedge b \tag{4.39}$$

(cf. [38] §4). For each relation $r \in R$, the disjuncts D_r in r is an object. Thus $D = \coprod_{r \in R} D_r$ is an object with a natural projection $\pi : D \to R$. Given maps $\lambda : R \to \mathcal{F}G, \rho : D \to \mathcal{F}G$, it is then possible to write each relation as

$$\bigwedge \lambda(r) \leq \bigvee_{\pi(d)} \bigwedge \rho(d). \tag{4.40}$$

This motivates the following definition.

[1] For instance, in **Sets**, we would need an object $\bigwedge_{i \in I} a_i$ for each set I. This collection is of course not a set

Definition 4.69 ([37] 5.1)**.** Let G, R and D be sets, and let $\lambda : R \to \mathcal{F}G, \rho : D \to \mathcal{F}G$ and $\pi : D \to R$ be functions. A *frame presentation* is a structure with the types D, R, G and the arrows π, λ, ρ.

Theorem 4.70 ([37] 5.4)**.** *Let X be a locale, and let $l : L \to X$ be the $\mathbf{Loc}/X$ representation of a locale internal to $\mathbf{Sh}(X)$ which is presented by the presentation P. If $f : Y \to X$ is a map, then the frame in $\mathbf{Sh}(Y)$ presented by $f^*(P)$ is given in the $\mathbf{Loc}/Y$ representaion as the space $l' : L' \to Y$ making the following square a pullback:*

$$\begin{array}{ccc} L' & \longrightarrow & L \\ \downarrow{\scriptstyle l'} & & \downarrow{\scriptstyle l} \\ Y & \xrightarrow{f} & X \end{array} \tag{4.41}$$

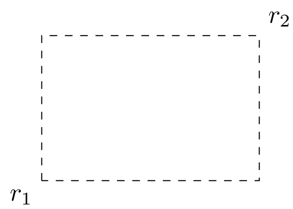
This theorem shows that if P behaves well with respect to geometric morphisms between sheaf topoi, then so does the frames presented by them.

An obvious corollary to the above theorem is that if L is the frame as presented in **Sets**, then the $\pi : L \times X \to X$ is the frame in $\mathbf{Sh}(X)$ given in the $\mathbf{Loc}/X$ representation.

4.7 The Complex Plane and the Upper Reals

Since the real and complex numbers are much used in standard quantum mechanics, it should come as no surprise that these concepts are important also in Topos Quantum Mechanics. Here, the corresponding constructions will be locales, the locale of complex numbers and the locale of upper reals.

Definition 4.71 (cf. [39] p. 441)**.** The *complex rationals* is the algebraically constructed set of numbers $r = p + qi$ where p and q are rational. The *locale of complex rationals*, denoted $\mathbb{C}_{\mathbb{Q}}$, is the locale of squares in the complex plane



(4.42)

identified with tuples (r_1, r_2), ordered by inclusion.

Definition 4.72 ([7] 2.4)**.** Given a topos $\mathcal{T}$. The internal frame of complex numbers $\mathbb{C}_{\mathcal{T}}$ is presented by the generators $(r, s) \in \mathbb{C}_{\mathbb{Q}}$, together with the following relations:

(i) $(r, s) \leq 0$ for $(r, s) \leq 0$.

(ii) $1 \leq \bigvee_{\text{all } r,s}(r, s)$

(iii) $(r, s) \leq (p, q) \vee (p', q')$ when $(r, s) \prec (p, q) \vee (p', q')$.

(iv) $(p, q) \wedge (p', q') \leq (r, s)$ when $(p, q) \wedge (p', q') \prec (r, s)$.

(v) $(r,s) = \bigvee_{(r',s') \prec (r,s)} (r',s')$.

Here. we remember the definition of rather below, $\prec$, from Definition 4.63.

Proposition 4.73 ([39] p. 442)**.** *The set generated by interpretation of this theory in* **Sets** *is the topology of* $\mathbb{C}$. □

Next, we consider the locale of upper reals.

Definition 4.74 ([5] §2.3)**.** The *locale of upper reals*, denoted $\mathbb{R}_l$, is the locale given by generators $(p,q) \in \mathbb{Q} \times \mathbb{Q}$ ordered by $(p,q) \leq (p',q')$ if and only if $p' \leq p$ and $q \leq q'$ with the relations:

(i) $(p,q) = \bigvee_{p<p'<q'<q} (p',q')$,

(ii) $(p_1,q_1) \wedge (p_2,q_2) = (\max(p_1,p_2), \min(q_1,q_2))$,

(iii) $\top = \bigvee_{p<q} (p,q)$.

To get an explicit description of these numbers, we will first need two definitions.

Definition 4.75 (cf. [29] D4.7.2 D4.7.3)**.** A function $X \to \mathbb{R}$ is called lower semicontious if it is continuous when $\mathbb{R}$ is given the topology generated by the sets $L_a = \{x | x > a\}$ for all $a \in \mathbb{R}$.

Definition 4.76 ([29] D4.7.3)**.** A lower semicontinuous function $X \to \mathbb{R}$ is called *locally bounded above* if for any $x \in X$, there exists an open neighbourhood V of x and a $q \in \mathbb{Q}$ such that $f(y) \leq q$ on V.

We note that it is possible to define the addition of two upper reals as pointwise addition.

Proposition 4.77 ([29] D4.7.3)**.** *For X a sober topological space, the object $\mathbb{R}_l$ internal to* $\mathbf{Sh}(X)$ *is given by the functor mapping V to the set of lower semicontinuous function $V \to \mathbb{R}$ which are locally bounded above.* □

4.8 The Alexandrov Topology

We have seen that sheaf categories allow different descriptions of internal locales, i.e. both as sheaves and as objects of $\mathbf{Loc}/X$. We will rely heavily on this when exploring covariant toy models in Chapter 6. However, our inspiration will not come from sheaves, but from covariant functors on a poset. Fortunately, covariant functors into **Sets** can be seen as sheaves, which will be shown in this chapter.

Proposition 4.78. *Let P be a poset. The collection of upwards closed sests, i.e. sets such that $x \in U, x \leq y$ implies $y \in U$, gives a topology on U. This is called the* Alexandrov topology on A.

Proof. A straightforward check of the axioms for a topology. □

Proposition 4.79. *The collection of sets $\mathcal{B} = \{\uparrow p\}_{p \in P}$ is a basis for $\mathcal{T}$.*

Proof. Clearly, each element of $\mathcal{B}$ is open. Furthermore, if $U \in \mathcal{T}$, then $U = \bigcup_{u \in U} \uparrow u$. □

Theorem 4.80. *Let P be a poset, and denote by $\hat{P}$ the poset seen as a space with the Alexandrov topology. Then there is an equivalence of categories*

$$\mathbf{Sets}^P \cong \mathbf{Sh}(\hat{P}). \tag{4.43}$$

Proof. We first note that the right hand side is a presheaf category $\mathbf{Sets}^{(P^{\mathrm{op}})^{\mathrm{op}}}$. We thus define $\tilde{P} = P^{\mathrm{op}}$, and will show the equivalence $\mathbf{Sets}^{\tilde{P}^{\mathrm{op}}} \cong \mathbf{Sh}(\hat{P})$.

Next, we define the functor $\mathrm{s} : \mathbf{Sets}^{\tilde{P}^{\mathrm{op}}} \to \mathbf{Sh}(\hat{P})$. By Proposition 4.29 and 4.79 it is enough to define the functor by showing how the image sheaf acts on the elements $\uparrow p$. Thus, we define $\mathrm{s}(F)(\uparrow p) = F(p)$ and send restrictions to restrictions. To see that this is a sheaf, we note that any covering of $\uparrow p$ must contain $\uparrow p$. But then, the sheaf requirement reduces to saying that any set of functions which is equal on all of $\uparrow p$ must come from a unique set of functions on $\uparrow p$, which is clearly true.

The other direction is a bit less obvious. We will assume that we are looking for an adjunction, and use to Yoneda lemma to define $\mathrm{c} : \mathbf{Sh}(\hat{P}) \to \mathbf{Sets}^{\tilde{P}^{\mathrm{op}}}$ by $\mathrm{c}(G)(p) = \hom(yp, \mathrm{c}(G)) = \hom(\mathrm{s}(yp), G)$. The latter expression is well defined, and is a functor by the naturality of the Yoneda lemma.

Now, we want to show that this definition gives an equivalence. It is clear that any sheaf G on $\hat{P}$ can be obtained by defining the presheaf $F(p) = G(\uparrow p)$, so the map is essentially surjective. Thus we need only show that

$$\hom(F_1, F_2) = \hom(\mathrm{s}(F_1), \mathrm{s}(F_2)) \tag{4.44}$$

But by 4.29, the arrows in $\mathbf{Sh}(\hat{P})$ are natural transformations if and only if

$$\begin{array}{ccc} s(F_1)(\uparrow p) & \longrightarrow & s(F_2)(\uparrow p) \\ \downarrow & & \downarrow \\ s(F_1)(q) & \longrightarrow & s(F_2)(q) \end{array} \tag{4.45}$$

commutes for each $\uparrow p$. This is clearly the same as asking that a diagram

$$\begin{array}{ccc} F_1(p) & \longrightarrow & F_2(p) \\ \downarrow & & \downarrow \\ F_1(q) & \longrightarrow & F_2(q) \end{array} \tag{4.46}$$

commutes, so the bijection is clear. □

Chapter 5

The Topos Approach

5.1 The Contravariant Approach

As mentioned at the outset, the main goal of this thesis will be to look at the topos approaches of quantum mechanics from a toy model perspective. Now, we have presented much of the theory of Topoi, and will thus move on to the description of the topos approaches of quantum mechanics.

We note that there are actually two slightly different topos approaches out there. The first one to appear is what is now called the contravariant approach, pioneered by Isham & Butterfield [1–4], and extended by Doering & Isham [41–44]. The second one is called the covariant approach, and was pioneered by Heunen *et al.* [5]. A comprehensive comparison of the two approaches can be found in [14].

Except as some historical background, the main importance for us of the contravariant approach is its similarity to the concept of measurement scenarios developed by [6]. These scenarios will be used to model toy models, and the main aim of Chapter 6 will be to incorporate these measurement scenarios into a setting more similar to the covariant approach. Thus this introduction will be rather brief. The interested reader is directed to the aforementioned articles and other texts on the subject, including [13, 45, 46].

The basic idea of the topos approach is that quantum theory behaves well when only commuting observables are taken into account. Remembering that each quantum system has associated to it a Hilbert space, the following definitions can be made.

Definition 5.1 ([45] 9.6)**.** Given a Hilbert space $\mathcal{H}$, consider the von Neumann algebra $\mathcal{B}(\mathcal{H})$ of bounded operators on $\mathcal{H}$. We let $\mathcal{V}(\mathcal{H})$ be the poset of abelian von Neumann subalgebras of $\mathcal{B}(\mathcal{H})$ ordered by inclusion. This will be called the *measurement scenario*, and each element $V \in \mathcal{V}(\mathcal{H})$ will be called a *measurement context*.

The terminology of measurement scenarios and measurement contexts are not used in [45], or in the Topos Approach literature in general. Instead, it comes from the Abramsky–Brandenburger approach which will be introduced in Section 5.2. We use it here to illustrate the similarities.

Now, both topos approaches depend on the Gelfand duality, which gives a dual correspondence between compact spaces and commutative C^*-algebras. It also has a constructive version in which compact spaces are replaced with compact completely regular locales.

Definition 5.2. A C^*-algebra A is an algebra over $\mathbb{C}$, together with an operation $(-)^*$ and a norm $|-|$ under which the space is complete. Furthermore, for $x, y \in A, \lambda \in \mathbb{C}$, the operations must satisfy:

(i) $(x+y)^* = x^* + y^*$,

(ii) $(xy)^* = y^* x^*$,

(iii) $(\lambda x)^* = \overline{\lambda} x^*$,

(iv) $(x^*)^* = x$,

(v) $||x^* x|| = ||x||^2$.

Theorem 5.3 ([47]). *The category of compact completely regular locales is equivalent to the dual category of commutative C^*-algebras, and this is true constructively. Furthermore, in the presence of the Axiom of Choice, compact completely regular locales are equivalent to compact Hausdorff spaces, giving a duality of the category of compact Hausdorff spaces and the category of commutative C^*-algebras.*

Proposition 5.4 ([45] §9.2). *In the classical case, the Gelfand spectrum of a commutative C^*-algebra C is given by the space of linear functionals on C which are multiplicative, i.e. for which $\lambda(AB) = \lambda(A)\lambda(B)$ for all $A, B \in C$, and has norm one, i.e. $\lambda(I) = 1$ where I is the identity of C.*

Equipped with this theorem, we are now ready to define the state space of the contravariant approach.

Definition 5.5. Given a Hilbert space with the corresponding poset (or measurement scenario) $\mathcal{V}(\mathcal{H})$, the *state space* is the functor $\Sigma : \mathcal{V}(\mathcal{H}) \to \mathbf{Sets}$ mapping each V to its Gelfand spectrum, and each inclusion to the restriction.

In order to get a good comparison with the toy models which we will look at later, we will look a bit more closely at finite-dimensional Hilbert spaces.

Proposition 5.6 (cf. [45] §9.1). *For $\dim(\mathcal{H}) < \infty$, any $V \in \mathcal{V}(\mathcal{H})$ is canonically generated by a set $\mathcal{P}_V$ of projection operators. Furthermore, if $V \subset V'$, then any projection operator $P \in \mathcal{P}_V$ is a linear combination of projection operators in the generating set of V'.*

Proof. From Lemma 2.4, we know that there is a basis $\mathcal{B}_V$ such that each b_i is an eigenvector for the operators of V. Now, group the basis elements by putting in the same group basis elements which have the same eigenvalue for all $A \in V$. Each group spans a space, and we take the collection $\{P_i\}$ of projection operators onto these spaces. It will be shown that these projections have the desired property.

Since each operator A operates as some constant a_i times the identity on each eigenspace, it is clear that $A = \sum_i a_i P_i$, i.e. that $\{P_i\}$ indeed spans V.

For the other direction, we will show that we can construct P_j using the operators of V. We know that for each $k \neq j$, there are two operators $A_k = \sum a_i P_i, B_k = \sum b_i P_i$ such that $a_k = b_k, a_j \neq b_j$. Consider $C_k = A_k - B_k$. Clearly $E_k \subseteq \ker C_k$ but $E_j \not\subseteq \ker C_k$. Thus scaling the operator $\prod_{k \neq j} C_k$ will give P_j.

The second statement of the theorem now follows easily. Indeed, if $V \subseteq V'$, then each maximal eigenspace E_i' of V' needs to be a subspace of one of the maximal eigenspaces of V'. Otherwise, it would be possible to construct a projection operator $P \notin V'$, using the construction above. The sum of the projections onto these spaces will be the desired operator. □

Proposition 5.7 (cf. [45] §9.2)**.** *For* $\dim(\mathcal{H}) < \infty$*, and* $V \in \mathcal{V}(\mathcal{H})$*,* $\Sigma(V)$ *is given by the set of linear functionals on* $\mathcal{H}$ *for which* $\{\lambda_i \mid \lambda_i(P_j) = \delta_{ij}\}$ *for all* $P_j \mathcal{P}_V$*.*

Proof. First note that for any projection operator P, and any $\lambda \in \Sigma_V$, it holds that

$$\lambda(P) = \lambda(PP) = \lambda(P)\lambda(P). \tag{5.1}$$

Thus $\lambda(P)$ is 1 or 0. Since the sum of all projections is I, it follows from linearity that

$$1 = \lambda(I) = \lambda(P_1 + \cdots + P_m) = \lambda(P_1) + \cdots + \lambda(P_m). \tag{5.2}$$

Combining these two equations shows that there is one P_i for which $\lambda(P_i) = 1$, and that $\lambda(P_j) = 0$ for all other P_j. □

We will finish this section by mentioning contextuality in the language of Isham and Butterfield. We know that each Hilbert space has associated a state space Σ in $\mathbf{Sets}^{\mathcal{V}(\mathcal{H})^{\mathrm{op}}}$. The corresponding state-space is said to be contextual if it has no global element in the topos $\mathbf{Sets}^{\mathcal{V}(\mathcal{H})^{\mathrm{op}}}$ ([1] §2.3.1).

5.2 Toy Models

In this section, we will introduce measurement scenarios. These were pioneered by Abramsky & Brandenburger [6], and has been further explored in a number of articles including, but certainly not limited to, [48–52].

We will start with the definition from [6]. These definitions will, however, be preliminary for our purposes, and the definition we will use is that of Definition 5.11. We thus present the constructions of Abramsky & Brandenburger mainly as background.

Definition 5.8 ([6] §2.2)**.** Let X and O be sets. Let the functor $\mathcal{E} : \mathcal{P}(X)^{\mathrm{op}} \to \mathbf{Sets}$ be defined by sending $U \in \mathcal{P}(X)$ to the set of functions from U to O, i.e. $U \mapsto O^U$, and sending $U \subseteq U'$ to the map $O^{U'} \to O^U$ given by restriction. This functor is called the O-valued and X-indexed *sheaf of events.*

In a measurement setting, the set X should be interpreted as a set of things that could be measured, and the set O as a set of measurement outcomes. A function $U \to O$ is then a way of, for each measurable, specifying an outcome. Given a specific situation in which a number of measurables are measured, it is clear that not every outcome need be equally probable, or even possible. To deal with this, we will look at R-distributions over the outcomes.

Definition 5.9 ([6] §2.3)**.** Let R be a semiring, i.e. a ring without the requirement that there is an additive inverse. An *R-distribution* on a set A is a function $d : A \to R$ with finite support, and with the property that

$$\sum_{a \in A} d(a) = 1. \tag{5.3}$$

The semirings that will be of most interest in the following is the semiring of nonnegative real numbers, $\mathbb{R}_{\geq 0}$, and the Booleans, $\mathbb{B}$. An $\mathbb{R}_{\geq 0}$-distribution is a usual finite probability distribution, specifying a probability for each element of X. A distribution over $\mathbb{B}$ is a possibilistic distribution, stating whether a certain outcome is at all possible, without reference to its probability.

Definition 5.10. Define the functor $\mathcal{D}_R : \mathbf{Sets} \to \mathbf{Sets}$ by mapping a set A to the set of R-distributions on A, and mapping a map $f : A \to B$ to the map $\mathcal{D}_R(f) : \mathcal{D}_R(A) \to \mathcal{D}_R(B)$ given by

$$\mathcal{D}_R(f)(d)(b) = \sum_{f(a)=b} d(a) \tag{5.4}$$

for $a \in A, b \in B$

Now consider the composite functor $\mathcal{D}_R\mathcal{E}$. For each set U of measurables in X, this will give the set of probability (or possibility) distributions on the set of functions O^U, i.e. all possible probability distributions for the different outcomes when the measurables in U are measured.

Now, let us introduce the definition of measurement scenario and empirical model that will be used in this thesis. We will use the following definition, heavily inspired by Abramsky's measurement scenarios and empirical models, but perhaps most resemble the topological models found in [52],

Definition 5.11. A *measurement scenario* M is a partially ordered set, An *empirical model* on M is a presheaf $F : M^{\text{op}} \to \mathbf{Sets}$ together with a topology on each $F(m)$.

Remark 5.12. Note that since this is a functor into **Sets**, we do not assume continuity of the restriction maps.

We think of each $m \in M$ as a measurement context. By a measurement context is meant a set of measurements being made on different parts of the system. Thus, the difference between a measurement context and a measurement (where we see the set of measurement as one measurement, giving a combined result) is only a question of semantics. We think of $F(m)$ as the set of possible outcomes for that measurement. The poset structure comes from inclusion of these sets. In other words, a measurement context m is smaller than n if the information obtained when doing the measurement corresponding to m is also obtained when doing the measurement corresponding to n.

We will often use the case with the discrete topology on each $F(m)$. In fact, whenever a specific empirical model is discussed, we will always assume $F(m)$ to be discrete unless something else is mentioned.

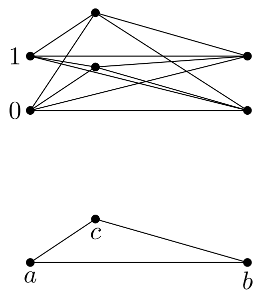
We note that this gives rise to a possibilistic scenario. We are only looking at what measurement outcomes are possible, and do not care about their probabilities. Probabilities for this type of measurement scenarios will be recovered later, when talking about states.

Definition 5.13. An empirical model is called *non-signaling* if for each $m \leq n$, the restriction map $F(\iota) : F(n) \to F(m)$ is surjective.

To see why the definition of non-signalling makes sense, we assume that we are in a situation where the observer Alice only has access to measurements in the context m, and that there are two more refined contexts $n_1, n_2 \geq m$ where the decision of another observer, Bob, decides which context we are in. Then, if there was $x \in F(m)$ which is only in the image of $F(n_1)$, then doing a measurement where the result x is obtained, Alice would be able to infer that we are in the context $F(n_1)$, and thus Bob has been able to signal to her which context we are in. In the rest of the text, we will assume all empirical models to be non-signalling.

It should be noted that there is a different way to look at non-signalling models. If we assume that Alice first makes her measurement, and obtained the value x, then Bob is forced to choose his measurement so that the context is n_1. Thus there seems to be a connection between non-signalling models and free choice of measurements. This has been explored further by Abramsky *et al.* [51].

For some measurement scenarios, it is possible to make graphical representations (cf. [48]). We will not make precise how such graphical representations can be made, but instead illustrate with an example. Let $M = \{a, b, c, ab, ac, bc\}$ where $a \leq ab, b \leq ab$ and so on. Further, let $F(a) = F(b) = F(c) = \{0, 1\}$ and let $F(ab) = F(a) \times F(b)$ and so on, with the obvious restriction maps. This can then be described by the following picture, where the line between a and b represent the context ab, and so on.



(5.5)

For many more examples of this kind of illustrations, see [53].

We note that the contravariant topos approach is a special case of a measurement scenario, with the empirical model given by the state space. More examples will be given in sections 6.2 and 6.3.

Also, the notion of Isham-Butterfield contextuality is transferrable to empirical models. We say that a model is contextual if the presheaf it defines has no global element. This is strong contextuality in the language of Abramsky & Brandenburger [6].

This notion of contextuality is actually stronger than the contextuality due to Spekkens (Definitions 2.8 and 2.9). Indeed, let us reinterpret the probability functions of Definition 2.8 to be functions into the Booleans instead, giving 1 for a possible outcome, and 0 for something impossible. Now, in a model which is non-contextual in the Spekkens sense, the possibility of a certain measurement outcome only depends on the initial state (and thus not on any other measurements). Thus it is possible to just one outcome for each measurement context. These measurements must work well together, because of the assumption that the measurements do only depend on the prepared state. This means that any model non-contextual in the sense of Spekkens must correspond to a model which is non-contextual in the Isham–Butterfield (or Abramsky–Brandenburger) sense. The converse is, in general, not true.

5.3 The Covariant Approach

We now turn to the covariant approach. The setting for the contravariant approach is the category of contravariant functors $\mathbf{Sets}^{C^{\mathrm{op}}}$ for a category of contexts C. Thus it might be no surprise that the setting for the covariant approach is $\mathbf{Sets}^C$, i.e. the setting of covariant functors into **Sets**.[1]

There is another difference between the two approaches, and that is what they use as their base categories. In the covariant approach, we do not require the underlying algebra to be von Neumann but settle with C^*. Then we have the following definition.

Definition 5.14 ([5] 4)**.** Let $\mathcal{A}$ be a C^*-algebra. Construct the category $\mathcal{C}(\mathcal{A})$ by taking as objects the commutative subalgebras of $\mathcal{A}$, and as arrows the inclusion. The *Bohrification of* $\mathcal{A}$, denoted A, is the tautological covariant functor from $\mathcal{C}(\mathcal{A})$ to **Sets** defined by $C \mapsto C$, with arrows mapping to inclusion.

In order to define a state space, we will use the Gelfand duality. We remember that its proof is constructive so that we can use it in arbitrary topoi, thus in particular in $\mathbf{Sets}^{\mathcal{C}(\mathcal{A})}$. Then we only need the following theorem.

Theorem 5.15 ([5] 5)**.** *The object A is a commutative C^*-algebra in $\mathbf{Sets}^{\mathcal{C}(\mathcal{A})}$ over the operations induced from $\mathcal{A}$.* □

Remark 5.16. Although the commutativity of the multiplication follows rather immediately, the fact that it is still a C^*-algebra is a bit less trivial. We note that it requires the construction of the internal object of complex numbers, which needs a natural numbers object. Furthermore, the proof of completeness for the norm uses dependent choice, which certainly does not hold for general topoi.

The above theorem makes it possible to use the internal Gelfand duality to associate to A a locale. This locale will be the state space for the contravariant approach.

Definition 5.17. Let $\mathcal{A}$ be a C^*-algebra, and A its Bohrification. Let $\mathcal{G}$ be the Gelfand functor from internal commutative C^*-algebras to internal compact completely regular locales. Then the *state space locale* of the quantum system associated with $\mathcal{A}$ is $\Sigma = \mathcal{G}(A)$.

It is possible to give an external description of the space Σ [54]. This description will be important later, so we will present it here.

Let $C \in \mathcal{C}(\mathcal{A})$, and let $\Sigma(C)$ denote the Gelfand spectrum of C. Then we can define the space

$$\overline{\Sigma} = \coprod_{C \in \mathcal{C}(\mathcal{A})} \Sigma(C) \tag{5.6}$$

and equip it with the following topology [54]. Let $U \in \mathcal{O}(\overline{\Sigma})$ if and only if the following two conditions hold for each $C \in \mathcal{C}(\mathcal{A})$:

(i) $U \cap \Sigma(C) \in \mathcal{O}(\Sigma(C))$,

(ii) For $C \leq D$ and $\lambda \in F(D)$, if $\lambda|_C \in U$, then $\lambda \in U$.

We then define

$$\overline{\Sigma}_U = \coprod_{C \in U} \Sigma(C) \tag{5.7}$$

[1] In its full generality, the covariant approach allows us to replace **Sets** as the image category by another category. However, we will not use that here.

with the topology inherited from $\overline{\Sigma}$. Then we have the following theorem

Theorem 5.18 ([54] 1)**.** *Given a C^*-algebra $\mathcal{A}$, the frame corresponding to the locale $\mathcal{G}(A)$ is given on $U \in \mathcal{A}$ by*

$$\mathcal{O}(\Sigma)(U) = \mathcal{O}(\overline{\Sigma}_U). \tag{5.8}$$

Besides the construction of a state space, we would also want to be able to talk about states in the covariant approach. We will here summarize the construction, with details provided in [5].

Definition 5.19 ([5] 9)**.** Given a C^*-algebra A, a map $\rho : A \to \mathbb{C}$ is called a *quasi-linear functional* if it is linear on each commutative subalgebra, and if $\rho(a + bi) = \rho(a) + i\rho(b)$ for self-adjoint a, b. If, for all a, $\rho(aa^*) \geq 0$, it is called positive. A positive quasi-linear functional on a unital C^*-algebra A is called a *quasi-state*.

Definition 5.20 ([5] 11)**.** Let X be a locale, and let $[0, 1]_l$ be the lower reals between 0 and 1. A monotone map $\mu : \mathcal{O}(X) \to [0, 1]_l$ that satisfies $\mu(U) + \mu(V) = \mu(U \wedge V) + \mu(U \vee V)$ and $\mu(\bigvee_i U_i) = \bigvee_i \mu(U_i)$ for a directed family U_i is called a *probability valuation*. Note that this definition is geometric.

Theorem 5.21 ([5] 14)**.** *For a unital C^*-algebra A, there is an isomorphism between the object of quasi-states and that of probability valuations on $\mathcal{G}(A)$.* □

Chapter 6

Covariant Toy Models

6.1 The construction

Now, we saw in section 5.3 how to construct the quantum state space locale as an object of **Loc**/X where X was the measurement scenario, equipped with the Alexandrov topology. Our idea is to study toy models using a copy of this construction.

Theorem 6.1 (cf. [14] §2.1)**.** *Let P be a poset (considered as a measurement scenario), and let $F : P^{op} \to \mathbf{Sets}$ be an empirical model, as defined in Definition 5.11. Let*

$$\overline{F} = \coprod_{p \in P} F(p). \tag{6.1}$$

Given the collection $\{\mathcal{T}_p\}$ of topologies on the different sets $F(p)$, define $\mathcal{T}$ to be the collection of subsets of $\overline{F}$ satisfying the following conditions:

(i) $U \cap F(p) \in \mathcal{T}_p$,

(ii) For $p \leq q$ and $x \in F(q)$, if $x|_p \in U$, then $x \in U$.

Then $\mathcal{T}$ is a topology on $\overline{F}$. Furthermore, if P is equipped with the Alexandrov topology and $\pi : \overline{F} \to P$ is the map sending all $x \in F(p)$ to p, then $\pi : \overline{F} \to P$ is a continuous map, and thus defines a locale internal to $\mathbf{Sh}(P)$.

Proof. Clearly the empty set and the whole space is in $\mathcal{T}$. Given $U, V \in \mathcal{T}$, we see that $(U \cap V) \cap F(p) = (U \cap F(p)) \cap (V \cap F(p))$ which is clearly in $\mathcal{T}$. Furthermore, if $x|_p \in U \cap V$, then x must be in both U and V, so $x \in U \cap V$, showing condition (ii). Moreover, for any family $\{U_i\}$, it is clear that the family $\{U_i \cap F(p)\}$ is open in $F(p)$, and if $f(x)$ is in some U_i, then x must also be in U_i. This shows that $\mathcal{T}$ is a topology. To conclude the argument, we note that $\pi^{-1}(\uparrow p) = \bigcup_{p \leq q} F(q)$ which is clearly an open set. □

Remark 6.2. If each $\mathcal{T}_p$ is discrete, this is just the Alexandrov topology on $\overline{F}$, where we have the ordering given for $\alpha \in F(p), \beta \in F(q)$ by $\alpha \leq \beta$ if and only if $p \leq q$ and $\beta|_p = \alpha$.

Definition 6.3. Let P and F be a poset (measurement scenario) and an empirical model. Then P equipped with the Alexandrov topology will again be called a *measurement scenario*. The internal locale defined by $\pi : \overline{F} \to P$ will be called a *covariant empirical model*.

Remark 6.4. Since any empirical model gives rise to a unique covariant empirical model, we will often drop the distinction, and let it be clear from context which one we are working with. We also want to note that, unless stated otherwise, that a sheaf on P will always mean a sheaf on P with the Alexandrov topology, whereas a presheaf on P will mean a presheaf on the ordered set P (i.e. not a presheaf on the topology of P).

The rest of the thesis will be devoted to exploring some implications and uses of this definition. As a first step, we give the explicit description of a covariant empirical model.

Proposition 6.5. *Let $\overline{F}$ be a covariant empirical model over P. Define the spaces $\overline{F}_U = \coprod_{p\in U} F(p)$ with the subset topology of $\overline{F}$. The internalization of $\pi : \overline{F} \to P$ as described in Definition 4.31, is the sheaf $\mathcal{O}\underline{F}$ defined on subsets by $U \mapsto \mathcal{O}(\overline{F}_U)$, and on subset relations by mapping $U \subseteq U'$ to the map $V \mapsto V \cap \overline{F}_{U'}$.*

Proof. We only need to note that the subobject classifier in $\mathbf{Sh}(\overline{F})$ is the functor sending each open V of $\overline{F}$ to the set of opens contained in V. Since for an open $U \in \mathcal{O}(P), \pi^{-1}(U) = \coprod_{p\in U} F(p)$, the proposition follows directly. ☐

Remark 6.6. By Remark 4.51, for $\iota : U \to X$, and $V \leq U$, we have $\iota^*(\mathcal{O}\underline{F})(U) = \mathcal{O}\underline{F}(U)$. Thus a straightforward comparison shows that $\iota(\mathcal{O}\underline{F})$ is the locale corresponding to $\pi : \coprod_{p\in U} F(p) \to U$.

Remark 6.7. We note that any contravariant state space can be seen as an empirical model. In this case, the locale gives exactly the locale described in Theorem 5.18. Thus, this construction gives a way to move from the contravariant to the covariant models.

We know that quantum mechanics is generally contextual, and thus that the state space, in general, has no global point. Similarly, our empirical models might not have global points. However, they do in fact have enough local points to make them spatial.

Lemma 6.8 ([29] C1.6.5)**.** *A locale map $f : Y \to X$ corresponds to an internally spatial locale in $\mathbf{Sh}(X)$ if and only if there is a Z and an epic locale map $g : Z \to Y$ such that $f \circ g$ is a local homeomorphism.* ☐

Theorem 6.9. *Any contravariant empirical model F over X is internally spatial.*

Proof. By Theorem 4.48, we only need to check the condition locally for any $p \in P$. Thus, we consider the open $\uparrow p$ together with the inclusion $\iota :\uparrow p \to X$. By Remark 6.6 the locale of interest is that represented by $\pi : \coprod_{p\in U} F(p) \to U$.

Define a *point of* F to be a set of elements $\{\lambda_q\}_{p\leq q}, \lambda_q \in F(q)$ with the property that for $q \leq q'$, $\lambda_{q'}|_q = \lambda_q$. Let S be the set of points of F with the discrete topology. The set $S\times \uparrow p$ is clearly étalé over p. Let $e : S\times \uparrow p \to \coprod_{p\leq q} F(q)$ by $\left(\{\lambda_q\}_{p\leq q}, q'\right) \mapsto \lambda_{q'}$. This map is a surjection, and thus an epimorphism in **Top**. But since the functor $\mathcal{O}$ is a left adjoint, this means that we also have an epimorphism in **Loc**. Since $\pi \circ e$ is étalé, it follows from Lemma 6.8 that the locale is spatial. ☐

Remark 6.10. Since a local point is a commutative diagram

$$\begin{array}{ccc} U & \longrightarrow & L \\ & \searrow\!\!\!\searrow & \downarrow \\ & & U \end{array} \tag{6.2}$$

we see that the points in $\uparrow p$ are exactly the points as mentioned in the Theorem, explaining the name. We note that a point is just a selection of outcomes in the different contexts which fits together.

Now, we know that the quantum state space comes from an internal commutative C^*-algebra. However, this is also true for many empirical models. In particular, all finite ones.

Definition 6.11 ([14] 2.5)**.** An empirical model $\overline{F} \to P$ is called finite if $F(p)$ is finite with the discrete topology for all $p \in P$.

We will now need some background from [14].

Definition 6.12 ([14] 2.5)**.** A continuous map $f : X \to Y$ between topological spaces is called *perfect* if the following holds:

(i) For any $y \in Y$, $f^{-1}(y)$ is compact,

(ii) For any closed C in X, $f(C)$ is closed (i.e. f is a closed map).

Lemma 6.13 ([14] 2.6)**.** *For a perfect map f, the locale $\mathcal{I}(f)$ is compact.* □

Lemma 6.14. *For any finite empirical model $\overline{F}$, the map $\pi : \overline{F} \to P$ is perfect.*

Proof. Since $\pi^{-1}(p) = F(p)$ is finite by assumption, it is also compact. Now let C be closed. We want to show that $\pi(C)^c$ is open. Let $a \in \pi(C)^c$, and let $a \leq b$. Given $\lambda \in \pi^{-1}(b)$, $\lambda|_a \in \pi^{-1}(a)$. Since C^c is open, it follows that $\lambda \in C^c$. Since this is true for each λ in the preimage of b, it follows that b is not in the image of C. Thus $b \in \pi(C)^c$, proving that $\pi(C)^c$ is open. □

The above shows that our empirical models give rise to compact locales. We thus only lack complete regularity for the Gelfand duality to apply.

Lemma 6.15 ([14] 2.9)**.** *Given a continuous map $f : X \to Y$, the locale $\mathcal{I}(f)$ is regular if and only if for any open $U \subset X$, and any point $y \in U$, there is an open $N \subset Y$ containing $f(y)$, and two opens $V, W \subset X$ with the properties $y \in V, V \cap W = \varnothing, f^{-1}(N) \subseteq U \cup W$.* □

Lemma 6.16. *The map $\pi : \overline{F} \to P$ has the property described above.*

Proof. Let U and $y \in U$ be given. Choose $N =\uparrow \pi(y)$. For each $\lambda \in \pi^{-1}(\pi(y))$, let $W_\lambda =\uparrow \lambda$. Set

$$V = \bigcup_{\lambda \in \pi^{-1}(\pi(y)) \cap U} W_\lambda, \qquad W = \bigcup_{\lambda \in \pi^{-1}(\pi(y)) \cap U^c} W_\lambda. \tag{6.3}$$

Now if $\alpha \in V \cap W$, then $r(\lambda) \in U \cap U^c$ which is clearly absurd. On the other hand, it is clear that $U \cup W$ covers $\pi^{-1}(N)$, since for all $\alpha \in \pi^{-1}(N)$ $\alpha|_{\pi(y)} \in U \cup U^c$. □

Now, by Theorem 4.80, the sheaf categories when using the Alexandrov topology behaves exactly like presheaf categories. Thus the following lemma is actually applicable.

Lemma 6.17 (cf. [35] §0)**.** *Assuming the axiom of countable choice, any compact regular frame is completely regular.*

Proof. Given $U \prec V$ we consider the set $\{W | U \prec W \prec V\}$. For a regular frame, this set is inhabited. Indeed, take the set $\{X | X \prec V\} \setminus \{Y | Y \prec U\}$. If this set is nonempty, then we are done. Otherwise, by regularity, we have $U = V$ in which case $V \prec V$, so V is in the set of interest. Given a choice axiom, it is also possible to choose one element from this set whenever required.

Now, it is straightforward to use dependent choice to choose sequences of the required form, given that $V \prec U$. Thus, in the presence of dependent choice, we have $V \prec U$ if and only if $V \ll U$, proving the statement. □

Theorem 6.18. *Any finite empirical model is a completely regular compact locale. Thus, by Gelfand duality (Theorem 5.3) any finite empirical model has an internal commutative C^*-algebra associated to it.*

Proof. By Theorem 4.15, the axiom of dependent choice holds in presheaf categories. By 4.80, the sheaf category of interest is equivalent to a presheaf category. Thus, this is a straightforward consequence of Lemmas 6.13 - 6.17. □

With this in mind, we will try to provide an explicit way to construct this internal C^* algebra. We remember that it is given by the internal maps from the locale, to the locale $\mathbb{C}$. From Proposition 4.73 and Theorem 4.70 we know that this locale is represented by $U \times \mathbb{C} \to U$, giving us the following theorem.

Theorem 6.19. *Let P be a poset (measurement scenario), and let F represent a finite empirical model. Let $U \subseteq P$ be open. Then any commutative diagram*

$$\begin{array}{ccc} F(U) & \xrightarrow{s} & U \times \mathbb{C} \\ & \searrow^{p} & \downarrow{\scriptstyle \pi_1} \\ & & U \end{array} \tag{6.4}$$

is given by mapping each connected component of $F(U)$ so that the composition $\pi_2 \circ s$ is constant. Here $\pi_2 : U \times \mathbb{C} \to \mathbb{C}$ is the projection. This states that $\mathcal{L}oc\left(S, \mathbb{C}_{\mathbf{Sh}(X)}\right)(U) = \mathbb{C}^{N(U)}$ where $N(U)$ is the set of components of $F(U)$.

Proof. Note that the set $\uparrow p$ for $p \in F(U)$ is an open set in the topology of $F(U)$. We first show that s needs to be constant on $\uparrow p$.

Indeed, assume that $q \in \uparrow p$, but that $\pi_2(s(q)) \neq \pi_2(s(p))$. Take an open $V \subseteq \mathbb{C}$ which contains $\pi_2(s(p))$ but not $\pi_2(s(q))$. Clearly $V \times \uparrow \pi(p)$ is an open of $\mathbb{C} \times U$. But since $s^{-1}(V \times \pi(p))$ contains p but not q, it can not be open, since all opens of $F(U)$ are upwards closed. It follows that s is not continuous.

Next assume that W is an arbitrary connected component of $F(U)$. Take a point $p \in F(U)$, and consider the open sets $A = \bigcup_{\pi_2(s(q))=\pi_2(s(p))} \uparrow q, B = \bigcup_{s(q)\neq s(p)} \uparrow q$. Since we know that the image of each $\uparrow q$ is constant, it follows that these two sets are disjoint. But since W is connected, it can not be decomposed into two nonempty disjoint open sets. Since $p \in A$, it follows that $B = \varnothing$. □

In Section 5.3 we introduced states for the covariant approach. We will use the isomorphism of 5.21 to define states for covariant empirical models.

Definition 6.20. Let X be a poset (measurement scenario), and let E be an empirical model, seen as a locale in $\mathbf{Sh}(X)$. Then a *state* in E is a probability valuation on E (as defined in Definition 5.20).

The following theorem gives an explicit description of states for finite empirical models.

Theorem 6.21. *Let F be a finite contravariant empirical model over a poset (measurement scenario) X (i.e. a functor $X^{op} \to$ **Sets** where each set is finite and equipped with the discrete topology) and let $\underline{F}$ be its covariant version. Then a global state of $\underline{F}$ is the object given by specifying for each $p \in X$ and each $\lambda \in F(p)$ a value $c_\lambda \geq 0$ obeying the restrictions*

(i) $\sum_{\lambda \in F(p)} c_\lambda = 1$ *for each* $p \in U$,

(ii) $\sum_{\lambda'|_p = \lambda} c_{\lambda'} = c_\lambda$ *for* $p \leq q, \lambda' \in F(q), \lambda \in F(p)$.

We then define the value on an arbitrary open $U \subseteq \underline{F}(X)$ by defining the natural transformation given at X by $\mu_X(U)(p) = \sum_{\lambda \in (U \cap (F(p))} c_\lambda$.

Proof. We first note that $\mu(U)$ is indeed a function $\mathcal{O}F(X) \to [0,1]_l$. Condition (i) gives that it takes values in the desired interval. We also see that for $p \leq q$, $\mu(U)(p) \leq \mu(U)(q)$ showing continuity. Monotinicity is obvious, since if $V \subseteq U$, then any λ in V is also in U. The conditions $\mu(U) + \mu(V) = \mu(U \wedge V) + \mu(U \vee V)$ and $\mu\left(\bigvee_i U_i\right) = \bigvee_i \mu(U_i)$ follows from the fact that they hold pointwise. Thus we know that a map on the above form is indeed a probability valuation.

We now note that any probability valuation need to take some value on each $\uparrow \lambda$, so we associated this number with c_λ. To make the function extend to a natural morphism, we need that for $\uparrow p \subseteq X$ the diagram

$$\begin{array}{ccc} \mathcal{O}F(X) & \xrightarrow{\mu_X} & [0,1]_l(X) \\ \downarrow & & \downarrow \\ \mathcal{O}F(\uparrow p) & \xrightarrow{\mu_{\uparrow p}} & [0,1]_l(\uparrow p) \end{array} \tag{6.5}$$

commutes. Clearly $\mu_{\uparrow p}(\mathcal{O}\underline{F}(\uparrow p)) = 1$ in order for μ to correspond to a probability valuation. From Proposition 6.5 we know that $\mathcal{O}\underline{F}(\uparrow p) = \mathcal{O}\coprod_{p \leq q} F(q) = \mathcal{O}\bigcup_{\lambda \in F(p)} \uparrow \lambda$. Thus we have

$$\sum_{\lambda \in F(p)} c_\lambda = \sum_{\lambda \in F(p)} \mu_X(\uparrow p) = \sum_{\lambda \in F(p)} \mu_{\uparrow p}(\uparrow p) = \mu_{\uparrow p}\left(\bigcup_{\lambda \in F(p)} \uparrow \lambda\right) = 1 \tag{6.6}$$

showing the first condition.

Next, we fix $p \leq q$. Let $U_{\lambda_i} = \bigcup_{\lambda|_p = \lambda_i} \uparrow \lambda$ for $\lambda_i \in F(p), \lambda \in F(q)$. Clearly $U_{\lambda_i} \subseteq \uparrow \lambda_i$ so by monotonicity, $\mu_X(U_{\lambda_i}) \leq \mu_X(\uparrow \lambda_i)$. But the converse follows from the fact that $\sum_i \mu(U_{\lambda_i}) = \sum_{\lambda \in F(q)} \mu(\uparrow \lambda) = 1 = \sum_{\lambda \in F(p)} \mu(\uparrow \lambda_i)$, showing the second condition.

Now, it follows from monotonicity that any μ_X must take the values $\mu(U)(p) \geq \sum_{\lambda \in U \cap F(p)} c_\lambda$, since $\bigcup_{\lambda \in U \cap F(p)} \uparrow \lambda \subseteq U$. Defining $V = \bigcup_{\lambda \in U^c \cap F(p)} \uparrow \lambda$ gives $\mu(V)(p) \geq \sum_{\lambda \in U^c \cup F(p)} c_\lambda$. However, since it must hold that $\mu(U)(p) + \mu(V)(p) = 1$, the two inequalities are indeed equalites. □

As described in 2.1, quantum mechanics distinguish between pure and mixed states. We remember that for pure states, there is a context where the outcome of a measurement is certain. We will use this intuition to define pure states in covariant empirical models.

Definition 6.22. Given a poset X, and an empirical model S, a state $\mu : S \to \mathbb{R}$ is called *pure at x* for $x \in X$ if there is an open $\uparrow \lambda \in \underline{S}(\uparrow x)$ such that $\mu(\lambda) = 1$. If there exists a maximal x such that μ is pure at x, μ is called *pure*.

Some measurement scenarios are "quantum" in the respect that it is possible to choose quantum measurements and outcomes as representatives for the outcomes in your empirical model. Thus, this gives a way to represent the empirical model using quantum mechanics. Formally, we make the following definition.

Definition 6.23. Let E be an empirical model over a poset X, and let Σ be a state space locale over $\mathcal{C}(\mathcal{A})$ as in definition 5.17. Then a *quantum realization* of E is a local homeomorphism $f : X \to \mathcal{C}(\mathcal{A})$ together with a locale arrow $f^*(\underline{\Sigma}) \to \underline{S}$. A measurement scenario where a quantum realization exists is called *quantum realizable*. Note that $f^*(\underline{\Sigma})$ is assured to be a locale because of the local homeomorphism requirement.

The map f thus picks out a quantum context for each context in X. Note that this map is not assumed to be injective, leaving us with the opportunity to assign the same quantum context to contexts in X, as long as the map is a local homeomorphism. Since we associate points of scenario locales with measured outcomes, the arrow $f^*(\underline{\Sigma}) \to \underline{S}$ associates a set of outcomes in the context $f(x)$ to represent it.

Since the set of quantum contexts is huge, even uncountable, and we are allowed to choose any contexts as our representatives, as long as they fit together properly, it does seem like a lot of finite measurement scenarios would be quantum realizable. However, not all are, and an example will be given in Section 6.3. In order to show this, we need a theorem which needs the following lemma.

Lemma 6.24. *If $\mu : X \to [0,1]_l$ is a probability valuation and $f : Y \to X$ is a frame map, then $\mu \circ f$ is also a probability valuation.*

Proof. We use constructive reasoning. Clearly $\mu(f(1)) = \mu(1) = 1$, and if $p \leq q$, then $f(p) \leq f(q)$ which in turn gives $\mu(f(p)) \leq \mu(f(q))$, showing monotonicity. Next, since μ is a state and since f preserves finite meets and arbitrary joins, we have $\mu(f(p)) + \mu(f(q)) = \mu(f(p) \vee f(q)) + \mu(f(p) \wedge f(q)) = \mu(f(p \vee q)) + \mu(f(p \wedge q))$ for all p and q, and $\mu\left(f\left(\bigvee_i p_i\right)\right) = \bigwedge_i \mu\left(f(p_i)\right)$. Thus $\mu \circ f$ satisfies all requirements for being a probability valuation. □

Remark 6.25. We should note that the corresponding locale map $l : X \to Y$ in many ways can be thought of as a random variable in the usual probability theoretic sense.

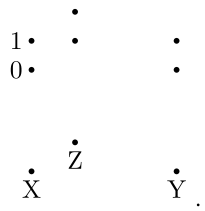
Theorem 6.26. *Any quantum realizable measurement scenario admits a state which is pure at x for each context x.*

Proof. We know that the quantum state space will allow a pure state at $f(x)$, just take the eigenstates of an operator in $f(x)$. Since the theory of valuations is geometric, there is thus a valuation on $f^*(\Sigma)$ mapping some $\uparrow \lambda$ to 1. Combining this with the frame map $\underline{S} \to f^*(\underline{\Sigma})$ gives the requied state. □

6.2 First example: Spekkens Toy Model

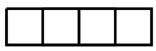
Spekkens toy model [8] is not a contextual theory in the Spekkens (or Isham-Butterfield) sense. However, we will see that it works very well with a contextual description. Thus it will provide a simple model to illustrate some of the concepts defined in this chapter. We start with an introduction to the model.

In the one "toy-bit" case, Spekkens toy model consists of three possible measurement contexts, here called the X, Y and Z contexts. There are no relations between these possible contexts. In each context, it is possible to get one out of two results, here called 0 and 1 for simplicity. If this was the whole story, our empirical model would have as base space the space $\mathcal{X} = \{X, Y, Z\}$, and $S(X) = S(Y) = S(Z) = \{0, 1\}$. Using the pictorial description presented in Section 5.2, we would get the rather boring picture



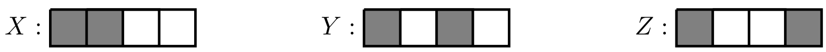
(6.7)

However, this is *not* the whole story. There is an assumption in the toy model that there are four so-called ontic states. Information that is inaccessible to the observer, but which still exists. These states also give information about the possible outcomes in the X, Y and Z measurements. Explicitly, we will represent the four ontic states as four boxes:



(6.8)

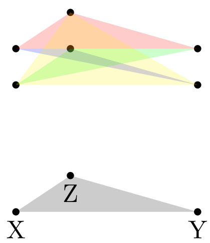
where we mark what state we are in by marking one of the boxes. The X, Y and Z measurements are now represented by the following pictures



(6.9)

where getting the result 1 is the same as getting a "yes" when asking if the state is in a grey box.

In order to accommodate this into our measurement scenario, we include another context called O.[1] Clearly there are now the relations $X \leq O, Y \leq O, Z \leq O$. We must also consider how the different ontic states restrict to the three subcontexts. We see that being in the first ontic state will give the result 1 in each of the different contexts, whereas the other four ontic states give 1 in one subcontext and 0 in the other two. Pictorially, this can be represented as



(6.10)

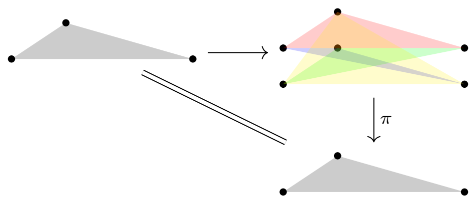
[1]This does not contradict the toy model if we agree that the measurement O, although existing, is impossible to make for a mortal observer.

where the grey triangle describes the context O.

It is now straightforward to find the associated C^*-algebra using Theorem 6.19 and the possible states using Theorem 6.21. Indeed, to present a state we need to assign a positive real to each ontic state in such a way that they sum to 1. Given such an assignment, the values assigned to each contextual state is determined by condition (ii) in Theorem 6.21. Thus a state is nothing else than a probability distribution over the ontic states.

Furthermore, the C^* algebra is $\mathbb{C}^4$ in the ontic context, and $\mathbb{C}^2$ in all of the subcontexts. (More explicit?)

We conclude this discussion by showing pictorially why Spekkens toy model is not contextual in the Isham–Butterfield sense. Indeed, let us see the diagram 6.10 as a space in $\mathbf{Loc}/\underline{S}$. We are then looking for maps making the following diagram commute:



(6.11)

But such maps clearly exists, as it is just a matter of mapping $\underline{S}$ to one of the coloured triangles.

6.3 Second example: The Popescu-Rohrlich box

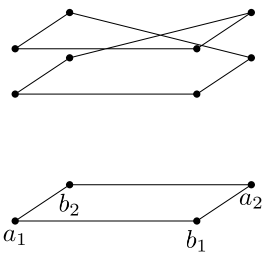
The next example is the Popescu-Rohrlich box [9], often shortened as the PR-box. This measurement scenario is originally thought of as two observers, let us call them Alice and Bob, both having access to two different measurements, for which 0 or 1 can be the measurement result. Let us call Alices measurements $\{a_1, a_2\}$ and Bobs measurements $\{b_1, b_2\}$. These are thus four of our contexts, but we also assume that it is possible for Alice and Bob to later meet and compare their measurements, which means that the contexts $a_i b_j$ for $i, j \in \{1, 2\}$ also exists. In each of the contexts a_1, b_i it is possible to get 1 or 0 as outcomes, and thus the outcomes in the contexts $a_1 b_i$ are ordered tuples of 0s and 1s.

Now, what makes the PR-box interesting is that we put some restrictions on what outcomes can be obtained together. Specifically, we only allow the combinations marked with 1 in Table 6.1.

	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
a_1b_1	1	0	0	1
a_1b_2	1	0	0	1
a_2b_1	1	0	0	1
a_2b_2	0	1	1	0

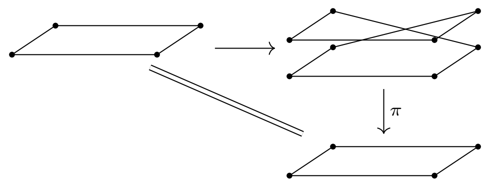
Table 6.1: Allowed outcomes in the PR-box measurement scenario

To see this as a presheaf, we make a poset P where $a_i, b_j \leq a_i b_j$, and a functor P^{op} where for instance $a_1 b_1 \mapsto \{(0,0),(1,1)\}$, and $a_1 \mapsto \{1,0\}$ with the obvious restriction maps. Seen pictorally, the PR-box is given as



(6.12)

We see immediately that the scenario is non-signalling. For each top-point, there is one line in each direction from that point. It should also be clear that there are no global sections for this box. Indeed, any attempt to find a section

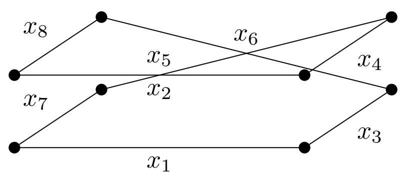


(6.13)

would require tearing the first rectangle, which would not make the map continuous.

Next, we want to consider the possible states of the PR-box. Although this situation can be handled in a more direct manner, we will introduce a general method for finding states. We note that a state is completely determined by its value on the opens $\uparrow x$ where x is maximal. Thus we associate to each such open a variable x_i. The conditions in Theorem 6.21 then gives relations between these variables, which gives a linear system of equations.

In this specific example, we have the variables $x_1, \cdots, x_8$, corresponding to each of the top-lines. We let the following picture illustrate our assignment of variables:



(6.14)

The conditions from Theorem 6.21 now gives for example the equations $x_1 + x_2 = 1$ and $x_1 = x_3$. Presented in matrix form, the full system of equations becomes

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{6.15}$$

We must also note that we are only interested in solutions where $x_i \geq 0$. The rank of the matrix can be calculated to be 8, so we expect a unique solution. Solving the system gives the single solution $x_i = 1/2$.

Saying the thing above using the terminology of states, we conclude that there is only one state, where each open $\uparrow \lambda$ maps to $1/2$. In particular, this shows that the PR-box admits no pure state. By Theorem 6.26, this gives a proof of the well-known fact that the PR-box is not quantum realizable. The original proof instead uses inequalities, more precisely the Tsirelson's bound [9].

References

1. Isham, C. J. & Butterfield, J. Topos Perspective on the Koschen-Specker Theorem: I. Quantum States as Generalized Valuations. *International Journal of Theoretical Physics,* 2669. ISSN: 0020-7748 (1998).
2. Butterfield, J. & Isham, C. J. A Topos Perspective on the Kochen-Specker Theorem. II. Conceptual Aspects and Classical Analogues. *International Journal of Theoretical Physics,* 827. ISSN: 0020-7748 (1999).
3. Hamilton, J., Isham, C. J. & Butterfield, J. Topos Perspective on the Kochen-Specker Theorem: III. Von Neumann Algebras as the Base Category. *International Journal of Theoretical Physics,* 1413. ISSN: 0020-7748 (2000).
4. Butterfield, J. & Isham, C. J. Topos Perspective on the Kochen-Specker Theorem IV. Interval Valuations. *International Journal of Theoretical Physics,* 613. ISSN: 0020-7748 (2002).
5. Heunen, C., Landsman, N. P. & Spitters, B. A Topos for Algebraic Quantum Theory. *Communications in Mathematical Physics* **291,** 63 –110. ISSN: 00103616 (n.d.).
6. Abramsky, S. & Brandenburger, A. The sheaf-theoretic structure of non-locality and contextuality. *New Journal of Physics* **13,** 113036 (Nov. 2011).
7. Banaschewski, B. & Mulvey, C. J. A globalisation of the Gelfand duality theorem. *Annals Of Pure And Applied Logic,* 62. ISSN: 0168-0072 (2006).
8. Spekkens, R. W. In defense of the epistemic view of quantum states: a toy theory. *Phys. Rev. A 75, 032110 (2007)* (Jan. 9, 2004).
9. Popescu, S. & Rohrlich, D. Quantum Nonlocality as an Axiom. *Foundations of Physics* **24,** 379–385. ISSN: 1572-9516 (1994).
10. A Einstein B Podolsky, N. R. Can Quantum-Mechanical Description of Physical Reality be Considered Complete. *Physical Review,* 777–780 (1935).
11. Bell, J. S. On the Problem of Hidden Variables in Quantum Mechanics. *Rev. Mod. Phys.* **38,** 447–452 (3 1966).
12. Kochen, S. & Specker, E. P. The Problem of Hidden Variables in Quantum Mechanics. *Journal of Mathematics and Mechanics* **17,** 59–87. ISSN: 00959057, 19435274 (1967).
13. Doering, A. Topos theory and 'neo-realist' quantum theory. *In Quantum Field Theory, Competitive Models, eds. B. Fauser, J. Tolksdorf, E. Zeidler, 25-47, Birkhaeuser (2009)* (Dec. 24, 2007).
14. Wolters, S. A Comparison of Two Topos-Theoretic Approaches to Quantum Theory. *Communications in Mathematical Physics* **317,** 3 –53. ISSN: 00103616 (2013).
15. Awodey, S. *Category Theory (Oxford Logic Guides)* ISBN: 978-0-19-923718-0 (Oxford University Press, 2010).
16. MacLane, S. & Moerdijk, I. *Sheaves in Geometry and Logic: A First Introduction to Topos Theory (Universitext)* ISBN: 978-0-387-97710-2 (Springer, 1994).
17. Bongaarts, P. *Quantum Theory: A Mathematical Approach* ISBN: 978-3-319-09561-5 (Springer, Cham, 2014).
18. Sakurai, J. & Napolitano, J. J. *Modern Quantum Mechanics: Pearson New International Edition* ISBN: 978-1-292-02410-3 (Pearson, 2013).

19. Friedman, A. *Foundations of Modern Analysis (Dover Books on Mathematics)* ISBN: 978-0-486-64062-4 (Dover Publications, 2010).
20. Roman, S. *Advanced Linear Algebra: 135 (Graduate Texts in Mathematics)* ISBN: 978-0-387-72831-5 (Springer, 2007).
21. Conrad, K. *The Minimal Polynomial and Some Applications* `http://www.math.uconn.edu/~kconrad/blurbs/linmultialg/minpolyandappns.pdf`.
22. Stenholm, S. & Suominen, K.-A. *Quantum Approach to Informatics* ISBN: 978-0-471-73610-3 (Wiley-Interscience, 2005).
23. Howard, M., Wallman, J., Veitch, V. & Emerson, J. Contextuality supplies the 'magic' for quantum computation. *Nature* **510,** 351 –355. ISSN: 00280836 (2014).
24. Spekkens, R. W. Contextuality for preparations, transformations, and unsharp measurements. *Physical Review A* **71,** 052108 (May 2005).
25. Aravind, P. Quantum Mysteries Revisited Again. *American Journal of Physics* (2004).
26. Peres, A. Two simple proofs of the Kochen-Specker theorem. *Journal of Physics A: Mathematical and General* **24,** L175 (1991).
27. Wikipedia contributors. *Quantum pseudo-telepathy — Wikipedia, The Free Encyclopedia* `https://en.wikipedia.org/w/index.php?title=Quantum_pseudo-telepathy&oldid=847579106`. [Online; accessed 3-August-2018]. 2018.
28. Cabello, A., Estebaranz, J. M. & García-Alcaine, G. Bell-Kochen-Specker theorem: A proof with 18 vectors. *Physics Letters A* **212,** 183–187 (Feb. 1996).
29. Johnstone, P. T. *Sketches of an Elephant: A Topos Theory Compendiumm vol. 1 (Oxford Logic Guides, 43)* ISBN: 0-19-853425-6 (Clarendon Press, 2002).
30. Borceux, F. *Handbook of Categorical Algebra: Volume 3, Sheaf Theory (Encyclopedia of Mathematics and its Applications)* ISBN: 978-0511525872 (Cambridge University Press, 2010).
34. Johnstone, P. T. *Stone Spaces (Cambridge Studies in Advanced Mathematics)* ISBN: 0521238935 (Cambridge University Press, 1982).
35. Baboolal, D. & Banaschewski, B. Compactification and local connectedness of frames. *Journal of Pure and Applied Algebra* **70,** 3–16. ISSN: 00224049 (1991).
37. Vickers, S. The double powerlocale and exponentiation: a case study in geometric logic. *Theory and Applications of Categories* **12,** 372–422 (2004).
38. Vickers, S. *Topology via Logic (Cambridge Tracts in Theoretical Computer Science)* ISBN: 0-521-36062-5 (Cambridge University Press, Cambridge, 1989).
39. Banaschewski, B. & Mulvey, C. The spectral theory of commutative C*-algebras: The constructive spectrum. *Quaestiones Mathematicae* **23,** 425–464. ISSN: 1727933X (2000).
41. Doering, A. & Isham, C. J. A Topos Foundation for Theories of Physics: I. Formal Languages for Physics. *J.Math.Phys.49:053515,2008* (Mar. 7, 2007).
42. Doering, A. & Isham, C. J. A Topos Foundation for Theories of Physics: II. Daseinisation and the Liberation of Quantum Theory. *J.Math.Phys.49:053516,2008* (Mar. 7, 2007).
43. Doering, A. & Isham, C. J. A Topos Foundation for Theories of Physics: III. The Representation of Physical Quantities With Arrows. *J.Math.Phys.49:053517,2008* (Mar. 7, 2007).
44. Doering, A. & Isham, C. J. A Topos Foundation for Theories of Physics: IV. Categories of Systems. *J.Math.Phys.49:053518,2008* (Mar. 7, 2007).

45. Flori, C. *A First Course in Topos Quantum Theory* ISBN: 978-3-642-35712-1 (Springer, 2013).

46. Flori, C. *A Second Course in Topos Quantum Theory (Lecture Notes in Physics)* ISBN: 978-3-319-71108-9 (Springer, 2018).

47. Coquand, T. & Spitters, B. Constructive Gelfand duality for C*-algebras. *Mathematical Proceedings of the Cambridge Philosophical society , Volume 147, Issue 02, September 2009, pp 323-337* (Aug. 11, 2008).

48. Abramsky, S., Mansfield, S. & Soares Barbosa, R. The Cohomology of Non-Locality and Contextuality. *ArXiv e-prints.* arXiv: `1111.3620 [quant-ph]` (Nov. 2011).

49. Abramsky, S., Soares Barbosa, R., Kishida, K., Lal, R. & Mansfield, S. Contextuality, Cohomology and Paradox. *ArXiv e-prints.* arXiv: `1502.03097 [quant-ph]` (Feb. 2015).

50. Abramsky, S., Barbosa, R. S. & Mansfield, S. Contextual Fraction as a Measure of Contextuality. *Phys. Rev. Lett.* **119,** 050504 (5 2017).

51. Abramsky, S., Brandenburger, A. & Savochkin, A. No-Signalling Is Equivalent To Free Choice of Measurements. *ArXiv e-prints.* arXiv: `1412.8523 [quant-ph]` (Dec. 2014).

52. Kishida, K. *Logic of Local Inference for Contextuality in Quantum Physics and Beyond* in *43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016)* (eds Chatzigiannakis, I., Mitzenmacher, M., Rabani, Y. & Sangiorgi, D.) **55** (Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2016), 113:1–113:14. ISBN: 978-3-95977-013-2.

53. Beer, K. & Osborne, T. J. Contextuality and bundle diagrams. arXiv: `1802.08424v1 [quant-ph]` (Feb. 23, 2018).

54. Heunen, C., Landsman, N. P., Wolters, S. & Spitters, B. The Gelfand Spectrum of a Noncommutative C*-algebra: a Topos-theoretic Approach. *Journal of the Australian Mathematical Society* **90,** 39 –52. ISSN: 14467887 (n.d.).