Quantum Field Theory: Motivating the Axiom of Microcausality

by

Jessey Wright

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Arts
in
Philosophy

Waterloo, Ontario, Canada, 2012

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Axiomatic quantum field theory is one approach to the project of merging the special theory of relativity with that of ordinary quantum mechanics. The project begins with the postulation of a set of axioms. Axioms should be motivated by reasonable physical principles in a way that illustrates how a given axiom is true. Motivations are often grounded in the principles of the parent theories: ordinary quantum mechanics or the theory of special relativity. Amongst the set of axioms first proposed by Haag and Kastler in 1963 is the axiom of microcausality. Microcausality requires the observables of regions at space-like separation to commute. This thesis seeks to answer the question'What principles from the special theory of relativity or ordinary quantum mechanics motivate, or justify, accepting microcausality as an axiom?' The first chapter will provide the necessary background to investigate this question and the second chapter will undertake that investigation. In conclusion, microcausality cannot be well-motivated by individual principles rooted in the special theory of relativity or ordinary quantum mechanics.

Acknowledgements

I am grateful to Doreen Fraser for her support and guidance throughout this project. I would also like to thank my readers, Dave DeVidi and Patricia Marino, for their insightful and helpful comments.

My fellow graduate students are also owed thanks, in particular Jim Tigwell and Lindsey Torma whose patience and discussions greatly contributed to the accessibility of the following arguments.

Contents

List of Figures

List of Abbreviations

AQFT = Algebraic Quantum Field Theory

GR = The General Theory of Relativity

OQM = Ordinary Quantum Mechanics

QFT = Quantum Field Theory

SR = The Special Theory of Relativity

Chapter 1

Quantum Field Theory in a nutshell

Quantum field theory is as much an attempt to carve a connection between two physical theories - the special theory of relativity and quantum mechanics - as it is an exercise in reclassifying a physical theory in terms of a new mathematical structure. Whether or not ordinary quantum mechanics and the special theory of relativity can 'peacefully coexist' has been a point of disagreement[1]. Quantum field theory aims to satisfy the physical requirements of both of these apparently conflicting theories. The first problem with such a project is finding a mathematical structure to build the theory upon. Ordinary quantum mechanics and the special theory of relativity are founded upon different structures, each lacking something important for describing the phenomena of the other.

Ordinary quantum mechanics utilizes the structure of the Hilbert space to describe phenomena. This structure accurately picks out the relationships between the various quantities of interest, the statistics of outcomes and relationships between states and operators. Ordinary quantum mechanics (OQM henceforth) describes the world in terms of interactions on a very tiny scale. The theory of general relativity is a field theory that teaches us about the universe at a large scale where objects can have subtle influences on the causal shape of space-time. If we want to unite OQM with GR then the most

[1] For a comprehensive overview of that debate see Maudlin, 1994.

natural step is to adapt the quantum formalism into a field theory so that it can account for the shape and structure of space-time. Unfortunately GR is a very complicated theory and it is unclear how one would go about making its predictions and those of OQM consistent within the same framework.

The theory of special relativity (SR henceforth) represents a special case of GR set in Minkowski space-time which is notable for being universally flat. A more reasonable task, that ideally would serve as a stepping stone to a generally relativistic quantum theory, would be to unite SR and OQM. The hope is that this will provide a good foundation for advancing the larger and more daunting project of creating a fully relativistic quantum theory.

One method of advancing the project of merging the special theory of relativity and ordinary quantum mechanics is through the study of axiomatic quantum field theory. The project begins with the postulation of a set of axioms. The axioms of quantum field theory examined in this thesis were presented originally as a plausible starting point for the project of axiomatizing quantum field theory [Haag and Kastler, 1963]. As axiom of a physical theory the story cannot end there and they should be appropriately justified by arguments from known physical principles. Ultimately, axioms of a physical theory should not be arbitrary, they should be true. A good motivation should establish that the proposed axiom reflects actual features of the world.

Often such motivations are grounded in the principles of the parent theories: ordinary quantum mechanics or the theory of special relativity. Amongst the set of axioms first proposed by Haag and Kastler in 1963 is the axiom now known as microcausality. Microcausality requires the observables of regions at space-like separation to commute. The goal of this thesis is to identify the motivation behind this axiom and in particular what aspect of the special theory of relativity or ordinary quantum mechanics provides that foundation. This first chapter will provide the necessary background to investigate this question and the second chapter will undertake that investigation.

I do not intend this chapter to be a comprehensive overview of these theories (OQM and SR) but instead brief, targeted summaries that get at the concepts key to the arguments at hand. The first section of this chapter will lay out the interpretive landscape and important definitions, the second will explicate the relevant features of the special theory of relativity and the third will do the same for ordinary quantum mechanics. The fourth section will introduce quantum field theory and the mathematical background needed to attend to the arguments that will follow. The final section of this chapter defines some of the axioms of quantum field theory, explicating what they are and briefly what they each try to accomplish.

1.1 Mathematics and Interpretations

When philosophically examining physics it is important to distinguish between implications rooted in the mathematical descriptions and interpretive descriptions of phenomena. I will make a careful effort to, where possible, make this transparent.

The mathematical formalism can be useful for informing possible sources of physical and interpretive tension as well as indicate unforeseen relationships between quantities. However, it is important to remember that the mathematics are, at their core, inspired by experimental results and known relationships between quantities. Any conclusions derived solely from mathematical principles must be backed up by sound, physically-based, reasoning.

Before this can be done one must apply an interpretation to the formalism. A vector, for instance, is a mathematical object and deciding what its physical analog is can be just as difficult as selecting vector spaces to model a particular system. Vectors, vector spaces and inner products are examples of what I am labelling as the mathematical formalism while position, space-time and measurement are examples of what I consider to be physical interpretations of the mathematics.

1.1.1 Operations and Operators

When discussing quantum mechanics the term "operator" is deployed in equal measure as a reference for a mathematical object and a physical process. In the interest of keeping the implications of the math separated as much as possible from the consequences of interpretation I will use the term 'operator' to refer to the mathematical object of the same name and the term 'operation' will be used to refer to its interpretive twin.

1.1.2 Observables

What constitutes an 'observable' is another interpretation-laden exercise that is necessary if one hopes to discuss quantum mechanics. While nomenclature is often indicative of the definition, in this case it can be misleading. Don't let yourself be too easily convinced that an observable is something which is observed (or that can be observed). Such notions can lead to privileging observation and measurement (the actions) when describing the universe.

Ruetsche classifies the observables of a theory as "the set of physical magnitudes that the theory recognizes" [Ruetsche, 8, 2011] and that is the definition I will be clinging to throughout. The mathematical form an observable takes within a given theory varies, but as far as I am concerned, an observable of a theory is a physical magnitude of the theory in question.

1.2 The Special Theory of Relativity

The majority of the motivations for the microcausality axiom in quantum field theory that will be examined have their roots in the special theory of relativity. Prohibitions against superluminal signalling (section 2.3), separability properties (section 2.3.1), the causal structure of space-time (section 2.4) and the independence of hyperplanes of simultaneity (section 2.5) are all motivations built upon principles of SR.

The special theory of relativity is a special case of GR that is set in Minkowski space-time. The curvature of Minkowski space is zero: it describes a space-time that is flat. While we know that space-time does not necessarily have zero curvature everywhere, GR shows that space-time is locally flat (has zero curvature in every small region of space-time). When working with the full theory of general relativity, for any system it is always possible to accelerate to a reference frame where locally (that is, in a small enough region around the system) the laws of physics behave like they do in the special theory of relativity. On a local scale, space-time looks like Minkowski space. This provides a rough and ready motivation for spending time uniting OQM with SR (instead of the more general and more successful GR). The hope is that if we can understand quantum phenomena in the context of special relativity than we will have, at least, a cursory understanding of quantum phenomena (locally) in the context of general relativity.

The fundamental postulate of the theory of special relativity has to do with the speed of light:

> **Law of Light**
>
> Every ray of light (in a vacuum) has the same speed, c, in all inertial frames[2] of reference. [Maudlin, 44, 1994]

No matter who you are, where you are going and which way you look, if you measure the speed of light you will find it to be c. The speed of light does not depend on the motion of the source nor the motion of the frame of reference from which the speed is measured [Maudlin, 44-45].

The Law of Light also requires the speed of light to be invariant under transformations between frames of reference. The law of light led to the discovery of counterintuitive phenomena that occur when an object travels at

[2] That the frame is an inertial one is an important qualification in the special theory of relativity that is eliminated in the general theory of relativity. For a more comprehensive discussion of the origins and meaning of inertial frames see: DiSalle, *Space and Time: Inertial Frames* [2009].

or very near to the speed of light. Well known examples of these phenomena are time dilation (time moves slower) and length contraction (distances are compressed).

1.2.1 Relativistic Space-Time

Newtonian physics takes place in Gallilean space, which is quite different from relativistic space-time. Perhaps the most obvious difference is the treatment of 'time' as a dimension (and hence the informative nomenclature - Gallilean space versus Minkowski space-*time*). Newtonian mechanics treats time as a parameter while the theory of special relativity recognizes that time is in fact a dimensional quantity. This difference is most visible in the geometric construction of Newtonian space, which uses 3-vectors (x,y,z), where Minkowski space-time invokes 4-vectors (t,x,y,z).

Gallilean space does not permit **any** velocity to be invariant. To relate quantities between different inertial frames in Newtonian mechanics we would use the Gallilean transformation. Since no velocity is invariant in Gallilean space, such a transformation will fail to preserve the constant speed of light. According to the law of light, the speed of light is invariant and so Gallilean space is not suited to describe relativistic space-time.

The Lorentz transform describes the relationship between measured quantities from the perspective of different inertial frames. As is required by the Law of Light, the speed of light remains fixed under the Lorentz transformation. When a quantity is invariant under a Lorentz transformation its value does not depend on the frame of reference it is measured (or observed) from. The speed of light is one such invariant quantity. The Lorentz transformation is used to relate reference frames in SR.

1.2.2 The Principle of Relativity

The principle of relativity is a vital postulate of the theory of special relativity. Harvey Brown, in a comprehensive study of the history and interpretations of the principle of relativity, argues that it can been expressed as a prohibition against frame dependence; all inertial frames are valid frames from which to observe and measure an event. Furthermore, the laws of physics, in order to be appropriately classified as laws, must be covariant under transformations [Brown, 2005]. The relationships between fundamental quantities must be preserved under coordinate transformations (translations, rotations, etc.). In the special theory of relativity the relevant coordinate transformation that laws must remain covariant under is the Lorentz transform. Brown argues that to describe the principle of relativity in terms of coordinate transformations is to make trivial the significance of the principle and brush aside the fact that the principle was not conceived of with transformations in mind [Brown, 36, 2005].

Roughly, the principle of relativity requires the relations between physical quantities to be the same whether the system is at rest or in uniform motion. Newton's statement of the principle of relativity is found in the *Principia* as Corollary V: "The motions of bodies included in a given space are the same among themselves, whether that space is at rest, or moves uniformly forwards in a right line without any circular motion." [Newton, Corollary V, Principia. From Brown, 35, 2005].

Newton's derivation of this principle assumes that accelerations are invariant and that forces and masses are velocity independent. These assumptions do not follow from Newton's postulated laws nor are they justified elsewhere in his work [Brown, 37-38, 2005]. While the principle itself is an important (and accepted) aspect of physical theories, it was not appropriately classified by Newton as a corollary.

Einstein, when he first presented the special theory of relativity, postulated the principle of relativity. Not only that, but he extended the scope of the principle to include electromagnetic phenomena [Brown, 2005]. Treating the principle as a postulate avoids the requirement for proof, the subsequent assumptions that Newton made in order to execute such a proof and makes clear the status of the principle. As a postulate behind the theory, an axiom if you will, the principle of relativity is taken to be a fundamental feature of the physical universe that cannot be derived from other principles or axioms.

It was reflection on the principle of relativity and a desire to generalize it further that led Einstein to develop the general theory of relativity (or at least to postulate general covariance) [DiSalle, 2009]. The principle, as postulated by Einstein, is: "If two coordinate systems are in uniform parallel translational motion relative to each other, the laws according to which the states of a physical system change do not depend on which of the two systems these changes are related to." [Einstein, 1905. From Brown, 74, 2005].

This has consequences for coordinate transformation laws, which the Lorentz transformation satisfies, and thus is often conflated with a covariance requirement. While covariance is one way to satisfy the principle of relativity, the principle itself is more general, and makes no explicit reference to, coordinate transformations and covariance[3].

1.2.3 The Relativity of Simultaneity

The order that two events in space-time occur in depends, in part, on the frame of reference that you observe them from. An event contained within the forward, Q_1 (or backward, Q_2) light cone of an event P will always occur after (or before) event P regardless of reference frame. Q_1 is said to be in the absolute future of P and Q_2 is said to be in the absolute past of P. The ordering of the events P and Q_1 will be invariant under Lorentz transformation. On the other hand, events outside of P's light cone, say R, are not invariant with respect to P. There will exist reference frames on which P and R occur simultaneously (the solid line frame in Figure 1), on which P occurs before R (the dashed line frame in Figure 1) and on which P occurs after R [Maudlin, 1994, 51-52]. When two events are simultaneous

[3] Maintaining the principle of relativity is important, at the very least, for ensuring that we can approximate isolation when performing experiments in a laboratory. Assuming that the principle of relativity is true allows experimenters to ignore a significant portion of the universe when trying to isolate the causes of experimental results. If the principle of relativity is false then it would be difficult to continue researching phenomena in the manner that we do because anything, anywhere could be a possible factor contributing to a given result. The principle of relativity, at least, lets us rule out events and objects at space-like separation. For more on the principle of relativity see Brown, 2005 and DiSalle, 2009.

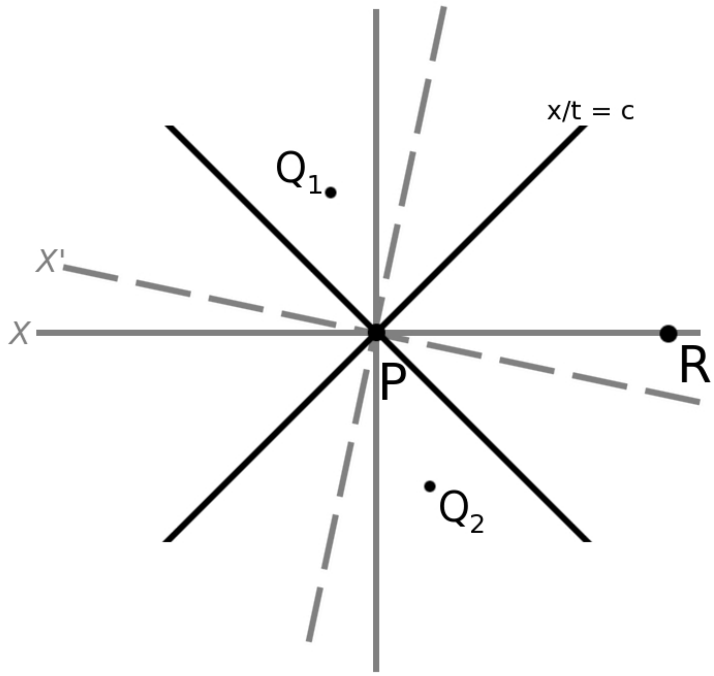

Figure 1.1: Space-Time Diagram: Solid line shows a frame of reference on which P and R are simultaneous, dashed line shows a frame on which P occurs before R. The bold lines have a slope of c and mark the forward and backwards light cones of P.

they are not contained within each other's light cones, hence simultaneity is relative.

Considering the principle of relativity in conjunction with the relativity of simultaneity it is natural to conclude that events at space-like separation must be independent. The principle of relativity tells us that there is no preferred frame of reference. The relativity of simultaneity implies that events which occur at space-like separation do not have a fixed ordering (the order depends on the inertial frame that you observe the events from). If events at space-like separation were able to influence one another then the principle of relativity would be violated because of the relativity of simultaneity. It would be possible to choose a reference frame such that the events occurred in any order and if the events can influence each other then their ordering matters. Thus, the final polarization state of the photons would depend on the frame of reference you observed the experiment with respect to; a direct violation of the principle of relativity.

This is not a conclusive argument since it has not been made clear in what sense space-like separated systems are prohibited from interacting nor the philosophical and physical limitations of these two principles. This will be addressed in greater detail in Section 2.3 when prohibitions against 'action at a distance' are evaluated as a possible motivation for microcausality and again in Section 2.5 when the relativity of simultaneity is specifically considered as a motivation for the axiom of microcausality.

1.3 Ordinary Quantum Mechanics

Consider ordinary quantum mechanics with the law of light and Lorentz transformations appended. This is one way philosophers have gone about exploring the relationship between SR and OQM. Maudlin shows that, for EPR states, different frames of reference have a non-zero probability of observing different outcomes of experiments. This shows that quantum phenomena can violate Lorentz covariance when one models the universe by naively attaching the principles of SR to OQM [Maudlin, 1994].

The goal of QFT is to uncover a way for SR and OQM to coexist. Maudlin's thought experiment shows that one cannot just take all the principles of SR and all of the principles of OQM and put them together to accomplish that goal. It is important to critically consider the implications of each of OQM and SR and understand how they relate each other, the structure of QFT and our broader understanding of the world.

1.3.1 Vectors, States and Operators

Quantum mechanics describes the world in terms of vectors and operators. These are mathematical concepts with accepted correlations to real objects (i.e.. states of electrons and photons are described by vectors and a measurement apparatus is described by an operator). The mathematics indicates how to calculate the vector state that will result from applying operators to vectors. This, in turn, correlates with performing a measurement on a photon and then observing the state of the photon after the measurement. What the mechanics do not indicate is what a 'measurement' (described by operators) or a 'wave function' (a linear combination of vectors which describes the properties of a quantum object) actually are and if these constructs have ontological content.

1.3.2 Non-locality

Assume that we have two photons which are maximally entangled, and we then take these photons and send them off in opposite directions and once they are sufficiently far away from each other (so far that a signal could not possibly travel from one to the other instantly) we perform a measurement on one. It turns out, because of the special way that the photons were prepared, that the outcomes of the measurements are correlated in a troubling way. A measurement carried out on one of the two photons can instantly change the wave function description of the other photon no matter how far apart they are [Albert, 1994].

Thanks to Bell's inequality and the experiments that followed its discovery and confirmed that quantum phenomena necessarily violate the inequality, we know that quantum states can exhibit apparently causal correlations in a way deemed by classical and relativistic intuitions to be 'non-local'. Whether or not these, Bell-type, correlations can be leveraged experimentally to actually violate laws of relativity is another question[4].

Non-locality will be relevant when discussing prohibitions against superluminal signalling as a possible motivation for microcausality.

[4] See Maudlin, 1994 for a closer examination of this problem

1.4 Quantum Field Theory

How we ought to study quantum field theory is an open question. In a recent exchange between Doreen Fraser and David Wallace the virtues of the two most common versions of QFT are debated. One version of QFT that is found in physics textbooks and used by physicists utilizes renormalization techniques to make calculations simpler. There are many variations of this class of QFTs utilizing cut offs and distinct renormalization techniques but they can be roughly categorized together as 'renormalized QFT' [Fraser, 2011]. It is this version of QFT that Wallace defends, primarily because it is the version of QFT utilized by practicing physicists [Wallace, 2011]. The other category of QFTs, and the one defended by Fraser, could be labelled 'axiomatic QFT'. It replaces renormalization techniques with rigorous axioms that are first postulated and then used to construct models [Fraser, 2011].

The axiomatic model has one major flaw, "to date, no model of the axioms has been constructed for any realistic interaction in four spacetime dimensions" [Fraser, 127, 2011]. This does not imply that no such model can be constructed, although some - such as Wallace - argue that if a model has not been constructed in the fifty or so years since the axiomatic program began then it is unlikely such a model will ever be constructed. And furthermore, renormalized QFTs are tied to actual models and indeed used by experimental physicists and so are already well ahead of the AQFT program on a practical level [Wallace 2011].

Whether or not AQFT is a better research program or one fated to fail is not my concern here. At the very least, since AQFT is better structured for carefully analyzing the foundations of quantum field theory, in part due to the rigorous axiomatic underpinning and careful attention to detail that goes into constructing models, that version of QFT will be utilized here.[5]

[5]The following section is a brief gloss over the relevant concepts and terminology that are deployed in or have bearing on later arguments. For a comprehensive introduction to the algebraic approach from a philosophical perspective see Ruetsche, 2011. For an introduction from a mathematical perspective see Halvorson, 2006.

1.4.1 Algebras and Vector Spaces

Algebras in a quantum field theory specify a set of *operators* over a space-time region while a *state* is a linear functional defined on a particular algebra that is both normed and positive. The algebra that the state is defined over constrains both the set of operators that can be used on the state (the elements of the algebra) and the space-time region over which it is appropriate to discuss that state (the region over which the algebra is defined). We shall use the term 'algebra' in a sense more restricted than the standard mathematical one. More precisely,

> **Algebra:** An algebra, $\mathcal{U}$, over the set of complex numbers, $\mathbb{C}$, is a set of elements which is:
>
> i. Closed under a commutative and associative operation of binary addition
>
> ii. Closed with respect to scalar multiplication by complex numbers
>
> iii. Closed with respect to a binary multiplication operation that is associative and distributive with respect to addition, but not necessarily commutative [Ruetsche, 74].

An important example of an algebra is $\mathcal{B}(\mathcal{H})$, the set of bounded operators on a Hilbert space.

> ***-algebra:** A *-algebra is an algebra, $\mathcal{U}$, that is closed under an involution, $* : \mathcal{U} \to \mathcal{U}$, satisfying:
>
> $$(A*)* = A, \; (A + B)* = A * +B*, \; (cA)* = \bar{c}A*, \\ (AB)* = A * B*$$

Where A and B are elements of the algebra $\mathcal{U}$ and c is a complex number.

$\mathcal{B}(\mathcal{H})$ is also a *-algebra [Ruetsche, 75].

C*-Algebra: $\mathcal{U}$ is called a C*-algebra if it is a self-adjoint subalgebra of $\mathcal{B}(\mathcal{H})$, for some $\mathcal{H}$, that is closed in $\mathcal{H}$'s uniform topology [Ruetsche, 77].

Uniform Convergence: A sequence, A_n, of operators on $\mathcal{H}$ converges to an operator A on $\mathcal{H}$ in the uniform operator topology if and only if $\| (A_n - A) \| \to 0$ as $n \to \infty$. Where $\| \text{o} \|$ is the Hilbert Space operator norm. [Ruetsche, 78].

Von Neumann (or W*) algebras are similar to C*-algebras except that they are closed in the strong operator topology instead of the uniform. That is a stronger closure condition, so W*-algebras have more operators than C*-algebras.

W*-Algebra: $\mathcal{M}$ is called a W*-algebra if it is a *-algebra of bounded operators on some Hilbert space $\mathcal{H}$, that is closed in $\mathcal{H}$'s strong operator topology [Ruetsche, 86].

Strong Convergence: A sequence A_n of operators on $\mathcal{H}$ converges to an operator A on $\mathcal{H}$ in the strong operator topology if and only if for each $| \psi > \in \mathcal{H} : | (A_n - A) | \psi > | \to 0$ as $n \to \infty$.

Commutant: For an algebra, $\mathcal{D}$, of bounded operators on a Hilbert space, its commutant $\mathcal{D}'$ is the set of all bounded operators on the Hilbert space that commute with every element of $\mathcal{D}$. If $\mathcal{D}$ is an algebra, then $\mathcal{D}'$ is too [Ruetsche, 86]

Because W*-algebras are closed in the strong (and not uniform) operator topology, it can be shown that a W*-algebra's double commutant is itself [Ruetsche, 87]. In fact, one can define W*-algebras by this property:

> **Double Commutant Theorem:** A von Neumann algebra M is a *-algebra of bounded operators such that $M = M''$ [Ruetsche, 88].

In addition to the above, the following definitions will also be relevant when separability conditions are on the table as possible motivations for microcausality:

> A linear map $T : A \rightarrow B$ is ***unit preserving*** if $T(I) = I$, where I is the identity of A.
>
> A linear map is ***positive*** when $X > 0$ entails that $T(X) > 0$.
>
> A linear map is ***completely positive*** when its natural extension, $T_n : M_n(A) \rightarrow M_n(B)$ is positive for every n, where $M_n(A)$ and $M_n(B)$ are the C* algebras of n-by-n matrices with entries from A and B respectively [Rédei, 1440, 2010].

1.4.2 States and Representations

States in algebraic quantum field theory are defined in terms of the observable algebra that they are defined on and so are not explicitly tied to a particular set of bounded operators on a Hilbert Space. This is in contrast to states in ordinary quantum mechanics which are also defined on the algebra of observables but are directly linked via the trace to density operators on the Hilbert Space. Furthermore, states in QFT are not required to be countably additive (unlike those in OQM) [Ruetsche, 90].

State: A state on a C*-algebra $\mathcal{U}$ is a linear functional $\omega : \mathcal{U} \to \mathbb{C}$ that is normed ($\omega(I) = 1$) and positive ($\omega(A * A) \geq 0$ for all $A \in \mathcal{U}$) [Ruetsche, 89].

Representations are morphisms into the algebra of bounded Hilbert space operators. C*-algebras admit such representations, and indeed any state on a C*-algebra can be used to generate a special class of representation that is particularly useful for doing QFT.

Representation (of a C*-algebra): A representation is a morphism $\pi : U \to \mathcal{B}(\mathcal{H})$ from the C*-algebra into the algebra $\mathcal{B}(\mathcal{H})$ of bounded operators on a Hilbert space $\mathcal{H}$ [Ruetsche, 83].

Additionally, a representation of a C*-algebra is itself a C*-algebra [Ruetsche, 86]. A recipe to translate from a state to a concrete Hilbert space representation and back would facilitate modelling complex quantum systems (such as systems which implement spontaneous symmetry breaking). Fortunately there is just such a thing, the GNS representation.

GNS Representation: For any state on a C*-algebra $\omega \in U$ there exists a Hilbert space, H_ω, representation $\pi_\omega : U \to H_\omega$ and a cyclic vector $|\xi_\omega\rangle \in H_\omega$ such that, for all $A \in U$, the expectation value the state ω assigns the algebraic element A is duplicated by the expectation value that the cyclic vector assigns to the Hilbert space operator $\pi(A)$. $(H_\omega, \pi_\omega, |\xi_\omega\rangle)$ is called the GNS representation [Ruetsche, 92].

1.5 The Axioms

There is more than one way to axiomatize quantum field theory. The approach adopted by this paper will be the same as the one used by Laura Ruetsche in *Interpreting Quantum Theories* [2011]. The following axioms were originally proposed by Haag and Kastler [1964] and later generalized by Dimcok [1980]. Ruetsche presents seven axioms (breaking one of the original axioms into two distinct axioms for the sake of clarity). For a more comprehensive analysis of these axioms see Ruetsche's book and the original paper by Haag and Kastler.

> **Isotony:** Where $\mathcal{O}_1$ and $\mathcal{O}_2$ are open bounded regions of $\mathcal{M}$, if $\mathcal{O}_1 \subset \mathcal{O}_2$ then $\mathcal{U}(\mathcal{O}_1)$ is a sub-algebra of $\mathcal{U}(\mathcal{O}_2)$ [Ruetsche, 105].

The isotony axiom requires algebras associated with regions of spacetime to preserve the inclusion relations of those regions. If one region is contained in another then the algebras that are associated with those regions are also so contained. Furthermore, it permits us to define a *quasi local algebra* $\mathcal{U}(\mathcal{M})$ which is an algebra associated with all of spacetime and sometimes called the *global state* [Ruetsche, 105].

> **Covariance:** Where Γ is the isometry group of $(\mathcal{M}, g_{ab})$, there is a group $\mathrm{G} = \{\alpha_\gamma,\ \gamma \in \Gamma\}$ of automorphisms of $\mathcal{U}(\mathcal{M})$ such that $\alpha_\gamma(\mathcal{U}(\mathcal{O})) = \mathcal{U}(\gamma\mathcal{O})$ for all $\mathcal{O} \subset \mathcal{M}$ and all $\gamma \in \Gamma$ [Ruetsche, 105].

The covariance axiom secures the Lorentz covariance requirement from the special theory of relativity. Whether or not this is sufficient to satisfy all of the requirements of the special theory of relativity remains to be seen (and will be discussed thoroughly in Chapter 2). For now it is sufficient to establish that indeed any quantum field theory which satisfies the covariance axiom also maintains Lorentz covariance in a way that would satisfy that requirement of SR.

> **Microcausality:** If $\mathcal{O}_1$ and $\mathcal{O}_2$ are spacelike separated, every element of $\mathcal{U}(\mathcal{O}_1)$ commutes with every element of $\mathcal{U}(\mathcal{O}_2)$ [Ruetsche, 106].

The second chapter is devoted to isolating an appropriate motivational stance to support this axiom. It will be useful to define spacelike separation:

> **Spacelike Separation:** Two regions, $\mathcal{O}_1$ and $\mathcal{O}_2$, are said to be spacelike separated when no point in $\mathcal{O}_1$ is connectable to any point in $\mathcal{O}_2$ by a causal curve. A causal curve is a curve whose tangent at any point is either time-like or light-like (but never spacelike) [Ruetsche, 106].

The vacuum axiom requires the existence of a vacuum state (or zero energy state).

> **Vacuum:** There exists a Lorentz-invariant state ω_o over the quasi local algebra $\mathcal{U}(\mathcal{M})$ whose GNS representation is faithful, irreducible and satisfies the spectrum condition [Ruetsche, 108].

The weak additivity axiom allows for the quasi local algebra that represents all of spacetime to be generated from the algebra over a local region by acting on that algebra with translations.

> **Weak Additivity:** For every closed, bounded region $\mathcal{O}$ of Minkowski spacetime $\mathcal{M}$, the closure in the C*-norm of the algebras $\mathcal{U}(\mathcal{O}+a)$ for $a \in \mathbb{R}^4$ coincides with the quasi local algebra $\mathcal{U}(\mathcal{M})$ [Ruetsche, 108].

Finally, the Primitive Causality axiom formalizes the notion that the past determines the future. It invokes the domain of dependence, which Haag and Kastler call "the causal shadow" [Haag and Kastler, 1964, 848].

> **Domain of Dependence (Causal Shadow):** The domain of dependence $D(\mathcal{O})$ of a space-time region $\mathcal{O} \subset (\mathcal{M}, g_{ab})$ is the set of points $p \in (\mathcal{M}, g_{ab})$ such that every inextendible causal curve through p intersects $\mathcal{O}$.

> **Primitive Causality:** If $\mathcal{O}_1 \subset D(\mathcal{O}_2)$, then $\mathcal{U}(\mathcal{O}_1) \subset \mathcal{U}(\mathcal{O}_2)$ [Ruetsche, 107].

The final axiom, which will not be utilized in the forthcoming arguments, is presented for completeness:

> **Primitivity:** There exists an irreducible, faithful representation of $\mathcal{U}_m$ [Ruetsche, 107].

It is worth noting that the axioms are logically independent, i.e., for each axiom there exists a model which violates only that axiom, and satisfies the rest [Ruetsche, 109].

With the interpretive landscape detailed, the important features of ordinary quantum mechanics and the special theory of relativity explained and the necessary tools in quantum field theory presented, the next chapter will begin the task of uncovering the motivation for the axiom of microcausality.

Chapter 2

Finding Motivation: The Foundation for Microcausality

The set of axioms originally postulated by Haag and Kastler [1964] includes one axiom that stands out as primarily concerned with the limitations of relativistic space-time. Indeed, Halvorson refers to it as "... the main *relativistic* assumption of AQFT" [Halvorson, 14, 2006]. That axiom was originally named "local commutivity" in Haag and Kastler's canonical paper and has also been referred to as "Einstein causality", "mutual commutivity" and, the term we shall use, "microcausality". The axiom follows:

> **Microcausality:** Let $\mathcal{U}_1$ and $\mathcal{U}_2$ be C*-algebras defined over space-like separated regions $\mathcal{O}_1$ and $\mathcal{O}_2$ respectively. If every element of $\mathcal{U}_1$ commutes with every element of $\mathcal{U}_2$, then $\mathcal{U}_1$ and $\mathcal{U}_2$ satisfy microcausality. Which is to say, if $A \in \mathcal{U}(\mathcal{O}_1)$ and $B \in \mathcal{U}(\mathcal{O}_2)$, then $AB = BA$.

The labelling of this axiom as relativistically rooted could be attributed to its use of the relativistic concept of space-like separation: a notion that only makes sense in the context of relativistic space time structure, a structure that is absent from ordinary quantum mechanics. The relativistic roots of microcausality do not end there. The requirement of mutual commutivity embodies, as will be further elaborated below, notions of compatibility, independence, separability and co-measurability. To put it very roughly, microcausality demands that space-like separated regions are 'causally shielded' from one another.

In ordinary quantum mechanics causal relationships are difficult to isolate. Maximally entangled states apparently conspire to produce correlated outcomes in situations where it is impossible for the states in question to communicate (either before or after the experiment). Explaining these causally deviant results has been a focus of interpretive and experimental efforts in ordinary quantum mechanics since their discovery. That microcausality imposes a limitation on causal behaviour that is defined in terms of space-like separation is good reason to suspect its ultimate motivation is rooted in relativistic requirements.

In the tradition surrounding quantum field theory, physicists and philosophers working in the field often set out with an operationalist mindset. Operationalism, as an interpretive stance, is a useful first step for isolating key intuitions in play and scratching the theoretical surface. For this reason, an understanding of the tenets of this stance will make a solid foundation for critically examining the ideas found at the core of quantum field theory.

Operationalism is an interpretive ideal that focusses its descriptions and definitions on measurements, operations and (in the case of quantum mechanics) outcomes and statistics. Roughly, operationalists focus their attention on the measurement process and treat it as the ultimate source for the definition of concepts [Chang, 2009]. Operational interpretations are natural when you try to attach meaning to physical terms, concepts and phenomena. As we will see, operational perspectives fail to make relevant distinctions between concepts that are closely related. That said, operationalist interpretations can be helpful when identifying the intuitions at play while interpreting a physical claim, which is a useful first step.

The operationalist insists that there is nothing ontologically interesting that is not accessible via measurement. The operationalist does not recognize properties which cannot be measured. An example of operationalist thinking can be found in Haag and Kastler's famous paper where they first formalized their axioms. They define a state and operation operationally: a state is "... a statistical ensemble of systems." and an operation is "... a physical apparatus which may act on the systems of an ensemble during a limited amount of time producing a transformation from an initial state to a final state." [Haag and Kastler, 1964]. These definitions are explicitly in terms of measurements, operations and statistical ensembles. If this is what states and operations are then physics can say nothing about that which has not (or cannot) be experimented upon. Furthermore, the focus on statistical ensembles makes it difficult to examine and make claims about individual properties possessed by specific particles or systems. Having an idea of what operationalism is we can now investigate what operationalism can tell us about microcausality.

Since an algebra is the set of operators as defined over a specified region of space-time, if the operations over one region of space-time commute with the operations over another, then the default interpretation is that the order that the operations are performed in does not matter. Of course, operations are not just performed out of context, they are performed *on* a particular state (at a particular time) and states are defined *over* algebras. When we say that the order of the operations does not matter, then we must mean that the order the operations are performed in on states does not matter to the statistics that describe the possible outcomes of performing the operations on those states. Microcausality seems to be telling us that operations performed in one region are *independent* of the operations performed in *any* other, space-like separated, region. This is the intuitive way to express what microcausality means physically and it is notably an operationalist interpretation. To fully understand what microcausality entails it is crucial to understand what it means for two operators or observables to commute.

In ordinary quantum mechanics, operators (which are self-adjoint elements of a Hilbert Space) represent observables. An observable is, roughly, a physical property (such as spin or polarization) that an object (such as an electron or photon) possesses. Commuting operators are said to describe compatible observables, which means that the observables can be measured and known simultaneously. Non-commuting observables are not compatible and so knowledge of one forbids knowledge of the other. Since the formalism that describes quantum mechanics is inherently stochastic, operationalist interpretations are fixed on the statistics of experiments (instead of particular outcomes). In operationalist terms, measuring one of a pair of non-commuting observables alters the statistics of a measurement on the other. If two observables do not commute then it matters in what order they are measured. Furthermore, two non-commuting observables cannot be simultaneously measured.

In quantum field theory, the commuting elements of an algebra are still operators but operators in QFT are different from operators in OQM, at the very least because in QFT operators are defined over a specific region of space-time. The operators in OQM are not the same mathematical objects as the operators in QFT, so care must be taken when using intuitions formed from understanding commutation relations in OQM in the context of QFT. Like commutability in OQM, commutability in QFT entails a kind of compatibility (more on compatibility in Section 2.4). In both OQM and QFT commuting operators are simultaneously realizable, but simultaneity has more intricate implications in the realm of QFT where relativistic considerations play a role (more on this in Section 2.5).

More concretely, Summers established that coexistence is equivalent to mutual commutivity [Summers, 203-204, 1990]. This result only holds for observable which can be written in the form of the projection operator, or basic observables. Two basic observables are said to coexist when they admit a decomposition which isolates the properties that the observables describe. To illustrate this, Summers uses the example of an apparatus with two binary displays. The observables linked to each display coexist if it is possible to write the two observables such that one represents the first display being ignored and the other represents the second display being ignored [Summers, 1990]. When the observables of interest are both basic and projections then their operators commute. Of course, this classification of coexistence is walking on the operationalist side of the fence, but (as is often the case) it provides a useful starting point and framing for our intuitions.

In what follows I will attempt to isolate the motivation behind postulating microcausality as an axiom. Section 2.1 begins this task by examining Lorentz covariance. In section 2.2 I will make explicit the relationship between microcausality and independence properties. Section 2.3 examines no-signalling theorems, the general prohibition against superluminal signalling and the property of separability. Section 2.4 considers the causal structure of space-time and in section 2.5 I will look at the principle of relativity and preferred hyperplanes. In section 2.6 I will consider a motivation that has its roots in ordinary quantum mechanics - the use of the tensor product structure to describe distinct systems. Finally, section 2.7 will concludes by reviewing this survey of motivations and attempt to settle whether the axiom of microcausality is appropriately motivated.

2.1 The Covariance Axiom

Since the Lorentz transformation can be used to derive phenomena such as time dilation, length contraction and the relativity of simultaneity it seems at first glance to be the primary feature of the special theory of relativity. If one assumes Lorentz covariance then the Law of Light follows immediately. Ruetsche argues that "... all the special theory of relativity can be taken to demand of a theory ... is that it exhibit Lorentz covariance. And this demand is met by QFTs in its scope, provided that they satisfy the Covariance Axiom" [Ruetsche, 109]

Axiomatic quantum field theory includes the **Covariance Axiom** which demands that all acceptable quantum field theories exhibit Lorentz covariance [Ruetsche, 109]. Ruetsche's argument, presented formally, follows:

P1: Quantum field theories which adhere to the covariance axiom (even if they violate the microcausality axiom) exhibit Lorentz covariance in the appropriate way.

P2: All that the special theory of relativity can demand of a theory is that it exhibits Lorentz covariance.

C: Therefore, the special theory of relativity cannot adequately motivate microcausality as an axiom. "Microcausality ... is to demand something more of [QFTs] than mere consistency with [the special theory of relativity]" [Ruetsche, 109].

Recall that the set of axioms considered in this paper are logically independent (see Chapter 1.5). Which is to say, for each axiom there is at least one model which violates only that axiom [Ruetsche, 109].

Quantum field theories which adhere to the covariance axiom satisfy Lorentz covariance in the appropriate way. The logical independence of the axioms and satisfactory treatment of Lorentz covariance by the covariance axiom secure the first premise in the argument above.

The second premise is not as well grounded. That Lorentz covariance is necessary for any theory that hopes to remain consistent with SR is certain, that it is all that is necessary is not. Implications of the special theory of relativity that inform our physical intuitions are not limited to the properties of Lorentz covariance. Only some principles of SR are directly derived from Lorentz covariance.

One could classify the implications of SR into two categories: those which follow from Lorentz covariance and those which do not. Amongst the set of those that do not are the principle of relativity and prohibitions against superluminal signalling (which is in part due to the principle of relativity). Superluminal signalling, the law of light and the relativity of simultaneity all have broader implications than the covariance axiom reaches. If adherence to Lorentz covariance is sufficient, as Ruetsche insists, to satisfy the demands of SR then appeals to the relativity of simultaneity, superluminal signalling and the law of light would be reducible to a requirement of Lorentz covariance. They are not.

The principle of relativity is just as vital a postulate to the special theory of relativity as the law of light. If not for the assumption of frame independence then the conclusions drawn from performing Lorentz transformations on hypothetical experimental set ups would not be possible. The principle of relativity makes it sensible to talk about observing events from different frames of reference and enables one to justify the use of the Lorentz transformation in the first place. The principle of relativity stands apart from Lorentz covariance, at the very least because it is an independent postulate of the theory and not derivable from other laws and axioms.

The law of light, while not logically independent from Lorentz covariance, still stands apart. While it is true that if one assumes the Lorentz transformation is correct that the constant speed of light immediately follows, the Law of Light is the fundamental postulate of the special theory of relativity and not a physical consequence of the requirement of Lorentz covariance. The form of the Lorentz transformation (which explicitly invokes the constant c) is *informed* by the Law of Light. Furthermore, if we look ahead to SRs successor, the Law of Light persists in the general theory of relativity while the Lorentz transformation is reduced to a special case. GR institutes general covariance in place of mere-Lorentz covariance. This alone is reason to privilege securing the Law of Light specifically, even in a system that already satisfies Lorentz covariance.

Superluminal signalling, the relativity of simultaneity and the law of light have something valuable to say about the way the universe works (or appears to work). As the next few sections will illustrate, even with the covariance axiom in place, there are many questions that arise within the context of QFTs that revolve around these properties and principles. If it is not already, it will become clear that Lorentz covariance is not "... all the special theory of relativity can be taken to demand of a theory set in Minkowski spacetime..." [Ruetsche, 109].

The one thing we can take away from this overview of the covariance axiom is that reliance on Lorentz covariance *alone* in order to motivate any other axiom of quantum field theory will fall flat. The Covariance axiom assures us that we have all that we could ever want to do with Lorentz covariance. Since the axioms are logically independent then the rest of them must stand on physical principles and requirements that are not Lorentz covariance. On the other hand, the special theory of relativity plays a larger role in describing the universe than Lorentz covariance captures. Relativistically based axioms beyond covariance are required to ensure that all of the physical demands of SR are appropriately satisfied by a given QFT.

2.2 Independence

Roughly speaking, operations which are independent are co-possible. Operationally, two observables are independent when they can be measured simultaneously. There are, at least, three categories of independence properties, two of which are clearly distinct from microcausality.

In Stephen Summers' survey of independence properties he defines three categories of independence. The first is *statistical independence* and will be defined as C* and W*-independence below. The second, *causal (or local) independence*, requires observables defined over space-like separated regions to be independent in some appropriate sense [Summers, 1990]. Finally, *kinematic independence* is a stronger formulation of local independence that is restricted to C*-algebras and requires that "... two observables which are represented by elements of a C*-algebra must commute" [Summers, 202]. Summers later describes local independence as the requirement of kinematic independence for spacelike separated observables [Summers, 1990].

Statistical independence captures the notion of preparation independence. Summers writes that, "... two quantum systems are statistically independent if each can be prepared in any state, how ever the other system has been prepared" [Summers, 202]. Kinematic independence requires two observables that are represented by elements of C*-algebras to commute. Operationally, this occurs when two observables are co-measurable. Note that statistical independence has to do with preparing states while kinematic independence has to do with measurement. It is not clear what the distinction between measurement processes and preparation processes amounts to, but as we will see these concepts are indeed distinct.

Summers defines local independence as: "...two observables associated with measurements in regions of spacetime that are spacelike separated from each other must be independent of each other in an appropriate sense, since no causal influences can propagate faster than the speed of light" [Summers, 1990, 202]. Microcausality is one way of enforcing this sense of independence, but the motivation here is based upon superluminal signalling. Summers also notes that the axiom of microcausality does not capture fully "... the physical notion of causal independence," [Summers, 202] and that the collection of physically motivated axioms for QFTs together result "... in a degree of independence for the local spacelike separated algebras of observables that goes well beyond [the] minimal requirement of kinematical independence" [Summers, 202][1].

This section will focus on statistical independence. In section 2.3 I will address superluminal signalling in detail. Kinematic independence will return in section 2.6 when I discuss a motivation from ordinary quantum mechanics.

Before discussing formalizations of independence a couple of notational definitions are in order:

> **Definition $\wedge$:** If e,f are projection operators on a Hilbert space $\mathcal{H}$, then $e \wedge f$ denote the projection onto the closed subspace $e(\mathcal{H}) \cap f(\mathcal{H})$ [Halvorson, 26, 2006].

> **Definition $\vee$:** If A, B are C*-subalgebras of some C*-algebra C, then $A \vee B$ denotes the C*-algebra generated by $A \cup B$ [Halvorson, 26, 2006].

Formally the three notions of independence that I would like to examine, in no particular order, are:

[1] In the interest of brevity, this paper will not address the axiom of microcausality in conjunction with the rest of the axioms. While there are interesting results to be explored, that may be relevant to motivating (or demotivating) microcausality; the goal here is to merely consider microcausality on its own.

Schlieder Property: Let R_1 and R_2 be von Neumann Algebras acting on a Hilbert space $\mathcal{H}$. The set $(R_1,\ R_2)$ satisfies the Schlieder Property if for all non-zero projections $e \in R_1$ and $f \in R_2$ then $e \wedge f \neq 0$ [Halvorson, 26, 2006].

If you can see how "$\wedge$" is analogous to conjunction in classical logic, then the Schlieder Property is analogous to logical independence [Halvorson, 26, 2006]. The Schlieder Property represents the 'logical' independence of two algebras in themselves and in particular the relations that hold between sets composed of elements of those two algebras.

The following formalizations of independence are specifically about the relationships between states. Already motivating microcausality via statistical independence is suspect, since microcausality has to do with operators and statistical independence has to do with states. What follows are formal definitions for Summers' statistical independence:

C*-Independence: Let A and B be C*-algebras. The set (A,B) is said to be C*-independent when, for any state $w_1 \in A$ and any state $w_2 \in B$, there exists a state $w \in A \vee B$ such that $w\ |_A = w_1$ and $w\ |_B = w_2$ [Halvorson, 26, 2006].

Halvorson immediately notes, after defining C*-independence, that it has an obvious operationalist motivation. This kind of independence embodies the idea that an observer in region A's choice to prepare a state cannot impact an observer in region B's ability to prepare a state [Halvorson, 26, 2006]. Halvorson also notes that C*-independence could alternatively be regarded as an explication of the independence of objects. Which is to say, two objects are independent only when any state of one is compatible with any possible state of the other. This is the case if they satisfy C*-independence [Halvorson, 27, 2006].

> **W*-Independence:** Let R_1 and R_2 be von Neumann algebras acting on a Hilbert space $\mathcal{H}$. The pair (R_1, R_2) is said to be W*-independent when, for every normal state $\psi_1 \in R_1$ and every normal state $\psi_2 \in R_2$, there exists a normal state $\psi \in R_1 \vee R_2$ such that $\psi\,|_{R_1}= \psi_1$ and $\psi\,|_{R_2}= \psi_2$ [Halvorson, 28, 2006].

A normal state on a von-Neumann algebra is an ultraweakly continuous state [Halvorson, 2006, 6].

Like C*-independence, operationally speaking, W*-independence expresses the notion that the preparation of normal states does not impact one's ability to prepare other normal states on another von Neumann algebra.

C* and W*-independence are identical in form but expressed relative to different algebraic structures. Recall that a representation of a C*-algebra is a morphism from that algebra to $\mathcal{B}(\mathcal{H})$, the set of bounded operators on a Hilbert space $\mathcal{H}$. One way to obtain a von Neumann algebra is to take a representation of a C*-algebra and close it in strong operator topology of its Hilbert space [Ruetsche, 86, 2011]. A von Neumann algebra is a *-algebra that satisfies a stronger closure property than a C*-algebra and so it contains more operators than a C*-algebra. The similarity of C* and W*-independence should come as no surprise since Summers classifies both as statistical independence [Summers, 1990].

Microcausality is closely related to the properties of statistical and logical independence. If any of these properties implied microcausality then the features of those properties would help illuminate the physical implications of microcausality. If microcausality implied any of the independence properties then, if the properties so implied are appropriately desirable (they accurately describe observed relations between physical systems) we might find motivation for the axiom here. Unfortunately implication does not link microcausality and these independence properties.

Halvorson shows that C*-independence does not imply microcausality [Halvorson, 27]. From Summers we know that if you assume local commutivity (*i.e.* microcausality) then W*-independence implies C*-independence.

Furthermore, when microcausality is assumed, C*-independence and the Schlieder Property imply each other [Summers, 1990]. Summers discussion of these properties and their relations reveals that W*-independence and microcausality are themselves independent. Likewise for C*-independence and the Schlieder Property. Microcausality greases the pathway connecting the formalizations of statistical independence, but is not implied by nor does it imply any of these three properties. Without microcausality the relationships connecting independence properties are not maintained. Microcausality does not entail nor is it entailed by any of these properties.

Here we may have found a possible motivation for postulating microcausality. When microcausality obtains W*-independence implies C*-independence and statistical independence and the Schlieder Property are equivalent. Without microcausality these relationships do not obtain. Thus, we ought to have microcausality as an axiom because it entails these connections between similar properties.

This fails to adequately motivate postulating microcausality as an axiom because appropriate motivations are not grounded in how we would like things to be, or how nice things work out when they obtain. A good motivation should establish that the proposed axiom accurately reflects how things actually are and should include some kind of physically based argument. That microcausality has nice (simplifying) implications is not enough to justify its status as an axiom. Such a motivation is essentially mathematical, and not appropriate for (solely) supporting a physical axiom. Additional evidence is required to show that the physical notions that these properties capture are actually related in the way that microcausality entails that they are.

At their foundation independence properties are a classification tool used to assess what kinds of relations hold between states and the measurements performed on arbitrary systems in arbitrary regions. Microcausality is a postulated rule which applies universally to all systems that fall under its scope. It is only natural for a law to impact categories, especially categories which are similar in scope to that law. If implications of an axiom fail to reflect reality then you have found grounds to challenge the postulation of the axiom. It requires more, however, than nice implications for an axiom to earn its stripes.

That microcausality implies logical relationships between statistical and logical independence properties is only grounds for accepting the axiom if those relationships obtain in reality. Until that is confirmed, that it implies those relationships merely expands the territory one might explore if you desire to reject the axiom. All you need to do is find a single physical example that violates the relationships that microcausality entails and the axiom would be eliminated (assuming, of course, that the independence properties in question access actual features of the world).

Returning to operationalism for a moment, statistical independence relations can be expressed in terms of limitations (or lack thereof) on state preparation. The operational interpretation of microcausality invokes the notion of co-measurability or co-realizability, not co-preparation. Microcausality is not in the business of securing independence of preparing states (perhaps obviously, since microcausality does not explicitly invoke states). On the other hand, while still wearing our operationalist hats, it is difficult to express the difference between a measurement and a preparation procedure. Operationally speaking they are the same; the process by which you measure a state and the process by which you prepare a state are identical. Here we find evidence that operationalist interpretations are not well suited for parsing the fine grained relationships between clearly distinct notions. It is obvious that there is a difference between independence of preparation and independence of measurement.

Microcausality is operationally interpreted as co-measurability and statistical independence is operationally interpreted as the freedom to prepare states in any order. However, the previous discussion showed that microcausality does not imply statistical independence and neither does statistical independence imply microcausality. Thus, independence of measurability does not imply independence of preparation and vise versa. The two processes must differ in a relevant way; we have both mathematical and physical reasons to distinguish these concepts. And yet, operationalist interpretations fail to make this distinction.

Notions of independence can only provide motivation for postulating microcausality if the relationships between them actually obtain in the world. Even if they are, that microcausality entails nice relationships between properties is not particularly sturdy ground to stand an axiom on. A better motivation would be more closely connected to the special theory of relativity or ordinary quantum mechanics. At the very least, we have found good reason to be suspicious of operationalist interpretations. In the next section I will examine the most commonly deployed motivation for postulating microcausality - that it prohibits superluminal signalling.

2.3 No-Signalling and Separability

Superluminal signalling refers to a class of (forbidden) phenomena: the process of sending a message or an interaction propagating at a speed greater than the speed of light. The most common motivation for microcausality makes an appeal to no-signalling results, which require the axiom of microcausality in order to be derived. Before deciding whether or not this will provide a good motivational foundation for the axiom of interest, it will be important to determine from where the prohibition against superluminal signalling arises. The obvious candidate is the special theory of relativity. After all, it does seem to have the greatest stake in the speed of light.

The first, and perhaps most important, thing to note about the special theory of relativity is that nowhere does it explicitly forbid anything from traveling at superluminal speeds. All SR has to say about the speed and light is that the speed of light is invariant, which is the content of the law of light. However, this provides plausible grounds for the claim. Examining the rate of change of the rest mass of a particle as it accelerates closer to the speed of light shows that as velocity approaches c then rest mass approaches infinity. This implies that in order to accelerate a particle beyond c one is required to produce an infinite amount of energy since mass and energy are proportionally related. Since producing an infinite amount of energy is physically impossible, it follows that no particle which is found traveling below the speed of light can ever accelerate to or beyond the speed of light [Maudlin, 65-71, 1994].

We are not out of the tunnel yet, as this only provides grounds for believing that c bounds the speed of the kinds of things that we are familiar with (electrons, protons, radiation and so on) but does not block the existence of objects which are born traveling at superluminal speeds. Particles that travel at speeds greater than the speed of light are called tachyons. They move through space (as we know it) like it was time and move through time (as we know it) like it was space. Which is to say, they are - if they exist - unlike anything we have or can experience. If tachyon's do exist the principle of relativity can be used to assure us that they cannot be the underlying cause of weird quantum correlations. If a tachyon appears to be responsible for such a correlation there will be another frame of reference from which the tachyon arrives too late to be causally responsible and yet another from which the tachyon could only be responsible if it was carrying relevant information before the experiment was even performed [Maudlin, 71-79, 1994].

We have evidence from SR that objects traveling at sub-luminal speeds cannot accelerate to or beyond the speed of light and that objects (which may or may not exist) traveling at superluminal speeds simply cannot be causally responsible (without invoking odd notions of backwards causation) for quantum correlations. And, just in case this isn't strong enough grounds to convince you, for the sake of argument, let's take it to be the case that superluminal signalling is something we want to prohibit.

Before returning our attention to the axiom of microcausality, the following definitions will prove useful:

> **Non-Selective Measurement:** A measurement is called non-selective when it does not select a particular state and instead produces a wave function expressed by a mixed state of all possible outcomes, each weighted by their respective Born Rule probability [Ruetsche, 109, 2011]. Mathematically, a non-selective measurement is a unit preserving, completely positive map [Rédei,

1440, 2010][2].

> **The Projection Postulate:** Let $\mathcal{A}$ be a von Neumann algebra (acting on $\mathcal{H}$). A self-adjoint operator Q that has a purely discrete spectrum λ_i and corresponding spectral projections P_i (in $\mathcal{A}$), is described by the projection postulate operation, T_{proj}, if it can be written as follows: $T_{proj}(X) = \sum_i P_i X P_i \in \mathcal{A}$ [Rédei, 2010].

A measurement is said to be represented by an operation in the projection postulate sense if it is possible to express the operator describing the measurement in the form of the projection postulate. Not all measurements can be so expressed; there are examples of interactions within quantum systems that cannot be described in the form of the projection postulate. A common example is when the observable in question does not have a discrete spectrum [Rédei, 2010].

If one assumes that the microcausality axiom holds then "... measurements represented by an operation in the projection postulate sense... carried out on a system localized in space-time region V_1 do not disturb a system localized in a space-like separated space-time region V_2" [Rédei, 1447, 2010]. The requirement of the projection postulate sense of measurement constrains the set of operations that this result applies to. Specifically, only to those measurements which can be written in the form of the projection postulate - which excludes selective measurements.

Microcausality is a necessary assumption for deriving no-signalling results such as the one above. The accompanying defence of microcausality proceeds as follows: microcausality has a justified place amongst the axioms of quantum field theory because it is sufficient for prohibiting superluminal signalling, which is forbidden by the special theory of relativity.

[2] Recall that a linear map $T : A \to B$ is unit preserving if $T(I) = I$, where I is the identity of A. And that a linear map is positive when $X > 0$ entails that $T(X) > 0$ [Rédei, 1440, 2010].

While it is not clear that superluminal signalling is explicitly prohibited by SR, there are good reasons to assume that it is (as elaborated upon above) and for the sake of argument we've assumed that it is. Ruetsche aptly notes that there are examples of physical systems which are excluded from no- signalling theorems. Microcausality implies no- signalling results that are restricted to non-selective measurements. The axiom does not prohibit selective measurements from engaging in causally bad behaviour. The class of correlation not constrained by microcausality are referred to (in the literature on Bell's inequality) as outcome dependent correlations. When two systems display outcome dependence the statistics of an experiment in one region depends on the outcome of a quantum measurement in the other. On one hand, the outcomes of quantum measurements are stochastic events and so the statistical dependence is beyond our control to manipulate, on the other hand causal relations are no less causal if we cannot manipulate them [Ruetsche, 111-112, and Maudlin 1994].

Even assuming that outcome dependence does not threaten 'causal good behaviour' in QFTs, no- signalling results entailed by microcausality are restricted in another way. Since the no- signalling theorem is restricted to measurements which are expressed in the projection postulate sense, they not only must be non-selective measurements but also must be measurements of *discrete* observables. It turns out that QFTs are full of observables which are accessible via non-selective operation but themselves are not discrete [Ruetsche, 112].

Microcausality is sufficient to rule out some classes of no- signalling, in particular, the class generated by non-selective measurements of discrete observables, but it fails to fully prohibit all possible cases of such action. If microcausality is to be motivated by its natural disposition to prohibit superluminal activity, and we are inclined to take action to prohibit such activity within our physical theory, the axiom will need to do more than block a few cases.

If this prohibition is the sole motivation for microcausality then explicitly prohibiting superluminal signalling would be a more effective way to secure the desired result since microcausality on its own is not such a prohibition, but only one of the many parts that come together to establish the desired result. No signalling theorems are one way of formalizing a prohibition against notions of action at a distance and if no signalling theorems are silent about the causal relationships between certain space-like separated systems then so too is microcausality, which is a prerequisite of those theorems. If we are interested in prohibiting action at a distance then microcausality has proven itself to be *insufficient* to do so universally.

2.3.1 Separability and Separatedness

Rédei and Valente sought to uncover the strongest no- signalling condition that microcausality does secure universally. As part of that project they defined and explored the properties of separability and separatedness. Separatedness is, essentially, a version of the no- signalling results just discussed. Separability is a weaker condition than separatedness and is related to notions of independence[3]. Separability refers to the ability to treat two systems as different systems.

The notions of separatedness and separability are invoked in order to investigate the question "... how local can local operations be in AQFT" [Rédei and Valente, 2010]. They first construct a typical no- signalling result (via operational separatedness), show that such a result is not universal (as was established for a similar result above) and then develop the notion of operational separability as a natural weakening of operational separatedness.

Operational *separateness* holds "if two subsystems of a larger quantum system are physically (causally) independent then an interaction with one of the subsystems that changes that subsystem's state, is restricted to the subsystem involved in the interaction and does not affect the other subsystem's state" [Rédei, 1441, 2010]. However, separateness is violated by a specific

[3] When the set of operations in question are compositionally closed and include the identity operator then C*-independence imply C*-separability. It is, however, not known if this relationship holds in general [Rédei, 2010].

class of operations defined on local algebras over space-like separated regions[4] [Rédei and Valente, 2010], which shows that operational separateness is too strong for quantum field theory. There are examples of non-separated local systems. Furthermore, the no- signalling result that holds for systems which are operationally separated only applies to Kraus operators. Since Kraus operators cannot be used to define a general operation, Rédei and Valente show that there exists *local* systems which violate operational separatedness and subsequently violate the entailed no- signalling result [Rédei and Valente, 2010].

In response to this discovery they weaken the condition of interest in order to explore the limits of locality within QFTs. The weaker separability condition (that is not violated in quantum field theories) is C*-separability:

> **(Operational) C*-Separability:** The pair $(\mathcal{A}_1, \mathcal{A}_2)$ of C*-subalgebras of C*-algebra $\mathcal{A}$ is called operationally C*-separable in $\mathcal{A}$ with respect to the sets of operations $(\mathcal{T}_1, \mathcal{T}_2, \mathcal{T})$ if every operation $T_1 \in \mathcal{T}_1$ has an extension $T \in \mathcal{T}$ which is the identity map on $\mathcal{A}_2$, and every $T_2 \in \mathcal{T}_2$ has an extension $T \in \mathcal{T}$ which is the identity map on $\mathcal{A}_1$ [Rédei, 1443, 2010].

The weakening comes in the form of a requirement that enforces 'causal good behaviour'. In this case, operations are not restricted to the region that they are local to but are such that their extension is the identity on other regions. The additional requirement of the identity extension blocks the existence of non-separable local systems that plagued separatedness. This is a weakening of operational separateness and so entails a weaker no- signalling result. However, unlike standard no- signalling results, assuming microcausality ensures that C*-separability holds in quantum field theories generally for regions which are spacelike separated [Rédei and Valente, 351, 2010]. On the other hand, just like other no- signalling results C*-separability still fails to

[4] In particular, operational separateness is violated by Accardi-Cecchini type operations defined on local algebras over space-like separated regions. See Rédei and Valente 2010 for more details.

block superluminal signalling in a category of relevant cases.

The method of motivating microcausality on the grounds of separability looks similar to the way no- signalling results are leveraged to the same end. Since the assumption of microcausality entails the that all space-like separated systems are C*-separable then, assuming C*-separability is a universal property space-like separated systems ought to possess, we should postulate microcausality. Phrased this way the motivation sidesteps the problems just discussed, namely the inability to prohibit superluminal activity via selective measurements. On the other hand, if microcausality entails that space-like separated regions are C*-separable, then it must be the case that such systems actually are C*-separable.

Even assuming that C*-separability is a property possessed by space-like separated systems, microcausality has no explicit stake in separability or superluminal causation. It has to do with the fundamental relationships between regions distinguished solely by their space-like separation. That postulating microcausality enables no- signalling results to be derived and properties such as C*-separability to be universalized are nice consequences, but as was argued above for the case of independence properties, nice consequences are not strong enough grounds to motivate an axiom. Such consequences provide space for disconfirmation, and for the axiom to stand the consequences of it (both mathematical and physical) must not fail to obtain in the world.

While it is interesting and informative to unravel the precise limits of the axiom's ability to prohibit superluminal signalling, that the separatedness conditions had to be restricted in such a way to enforce causally good behaviour is further evidence that microcausality could do better in the realm of no- signalling results. As Ruetsche neatly puts it, "... there is room to worry that Microcausality would be better motivated if it were demonstrably sufficient to rule out a wider variety of signalling than it at present is." [Ruetsche, 112]. The separability results still do not solve issues of outcome dependence (and selective measurement) nor do they extend the result to cover discrete operations. In fact, the C*-separability results is a weakening, not an extension, of the insufficient motivation for microcausality that is no-signalling.

Having established the role of microcausality relative to the theory of special relativity in general, set up the hierarchy of some common independence properties and grappled with the most prevalent motivation for the axiom, we will change gears in the next section and explore the axiom's origins in search of a stronger foundation.

2.4 Causal Structure, The Original Motivation

Haag originally motivated the axiom as follows: "The main principle expressed by it is the causal structure of events. Two observables associated with space-like separated regions are **compatible**. The measurement of one does not disturb the measurement of the other. The operators representing these observables must commute." [Haag, 107, 1992. Emphasis mine].

Compatibility is essentially rooted in notions of disturbance. Two systems can be said to be compatible if actions performed on or by one system do not disturb the other. This notion could be identified as an independence or separability property. Operationally, compatibility refers to comeasurability. Two systems are compatible (in the operationalist sense) if they are simultaneously measurable [Summers, 1990]. Given that Haag (and Kastler) invoke operationalist accounts of states and observables it is probable that their motivation for microcausality is operationally rooted.

Ruetsche attempts to develop a formal motivation on Haag's behalf. Let A and B represent arbitrary operations in space-like separated regions O_1 and O_2 respectively and $|\psi\rangle$ represent a global state (which is an element of the quasi-local algebra $\mathcal{U}(\mathcal{M})$).

P1: If $AB\,|\psi\rangle \neq BA\,|\psi\rangle$, then the form of the global state after performing operations in O_1 and O_2 would depend on the order that those operations were performed in.

P2: Since the regions are space-like separated, we know that there is no matter of fact about the order of the operations.

C: Therefore A and B must commute [Ruetsche, 110, 2010].

Ruetsche then proceeds to defeat this reconstruction by challenging the notion of the global algebra admitting a state. That notion is in conflict with the motivation behind the Vacuum Axiom; the idea of a measurement in ordinary quantum mechanics does not accommodate allowing the quasi-local algebra $\mathcal{U}(\mathcal{M})$ to admit a state. "If $|\psi\rangle$ is [a] global state, it's a state for all space and all time, and the notion of changing it by an operation executed within spacetime is nonsense." [Ruetsche, 110]. Then $|\psi\rangle$ must not be a state for all space and all time, so instead Ruetsche casts it as a state *at a time*.

If $|\psi\rangle$ is a state at a time then it is defined on a particular hyperplane of simultaneity Σ. In which case $|\psi\rangle$ is defined on the algebra $\mathcal{U}(\Sigma)$. The same problem with $|\psi\rangle$ as a global state appears when one introduces the Primitive Causality axiom[5]. "Primitive Causality implies that the algebra for all of spacetime $\mathcal{U}(\mathcal{M})$ coincides with the algebra $\mathcal{U}(\Sigma)$. Any state on one is automatically a state on the other, and so not subject to alternation by activity within the spacetime for which it's a state." [Ruetsche, 110]. That this reconstruction is disabled so quickly is one indication of its weakness. Ruetsche has built a straw-man to defend Haag.

The use of a global state to capture the notion of compatibility was ill fated from the outset. Haag did not have in mind the idea of measurements on global states when he first presented the axiom. When first presenting the principle of locality Haag and Kastler state that, "... observables in causally disjoint regions are always compatible" [Haag and Kastler, 1964]. The focal point of the principle is that observables in *causally disjoint regions*, which are not represented by a global state (neither in ordinary quantum mechanics nor in quantum field theories). Ruetsche was spot on when she centred her formal reconstruction of Haag's position in terms of *local* operations,

[5]Recall that Primitive Causality, defined in Section 1.5, essentially requires the future events in a region to be determined solely by the events contained in its 'causal shadow'.

but ignoring "... more nuanced and subtle models of local operations..." [Ruetsche, 111] was an error. The notion of local operation that Haag had in mind is, at least, more nuanced than a global state or a global state at a time since both of these notions invoke a global state.

Reutsche's construction does have an upside. It clearly establishes that, in the absence of a more nuanced sense of local operation, microcausality cannot claim to be necessary to rule out action-at-a-distance [Ruetsche, 111]. This would open the door to discuss superluminal signalling (since microcausality *can* claim to be sufficient for such a prohibition), but we've already ruled superluminal signalling out as a motivation in the previous section.

Haag's motivation for microcausality rests upon a requirement of compatibility between spacelike separated systems. Haag writes that if two states are compatible then "[t]he measurement of one does not disturb the measurement of the other" [Haag, 107], and that postulating microcausality implies that the best way to achieve that requirement is to insist upon mutual commutivity between operations at space-like separation.

Mutual commutivity is not the only way to ensure that observables in space-like separated regions are compatible with each other. Separability and independence are examples of other formalized properties which perform a similar role. If compatibility between space-like separated regions is a physical requirement, then postulating microcausality to do so needs to be motivated. At the very least, one needs to provide reasons to adopt microcausality as an axiom instead of independence or separability, all of which can be used to secure compatibility between space-like separated regions.

C* and W*-independence both express the notion that independent states are co-measurable: they can both be measured simultaneously. Unfortunately, co-measurability and non-disturbance of measurement (which Haag's compatibility requires) are not *exactly* the same properties. Furthermore, independence is expressed in terms of states and not measurement operators and so interpretive work would need to be done to explicitly establish how independence takes the shape of a compatibility requirement. Separability, on the other hand, is closer to Haag's definition of compatibility as it deals explicitly with operators and the ability to perform operations on states without having an impact on states in a different region.

Previously we explored these notions and found them lacking as motivation for microcausality, but those discussions do no prohibit these properties from being postulated as a replacement for microcausality. In order to appropriately universalize them they would need to be granted the same scope as the axiom - universal application to spacelike separated regions. Once this extension is applied each of these positions arrives at the same place we find microcausality in now, seeking adequate physical motivation for being postulated as an 'axiom'. Microcausality must have some inherent value to have been initially postulated when these other properties were not, and to have survived as a member of this particular axiomatic program for the last fifty years.

If Haag's requirement for compatibility can be essentially described as a requirement for independence or separability then the results of section 2.2 and 2.3 can be applied and the same conclusion follows: microcausality can't find an adequate motivation here. There must be something deeper underwriting the axiom.

In the canonical paper where Haag and Kastler present their axioms for QFT they open with a definition of what they take to be the principle of locality:

> "The essential feature which distinguishes quantum field theory within the frame of general quantum physics is the principle of locality. This principle states that it is meaningful to talk of observables which can be measured in a specific space-time region and that observables in causally disjoint regions are always compatible." [Haag and Kastler, 1964].

This statement of the principle of locality contains two claims. The first is about operations in specific spacetime regions – which implicitly assumes that operations defined on spacetime regions can be performed in those spacetime regions. The second, if we assume that spacelike separated regions are causally disjoint[6], is a statement of microcausality. Haags' local commutivity axiom (which is our microcausality axiom) is embedded within the principle of locality. And, the principle of locality is the motivating assumption behind Haag and Kastler's project. It seems like Haag and Kastler have the axiom of microcausality already in place. It's postulation goes without motivation because it is part of the background, buried in the description of "[t]he essential feature which distinguishing quantum field theory..." [Haag and Kastler, 1964].

Recall that the axiom of microcausality is motivated, according to Haag, because algebras over space-like separated regions need to be compatible. The principle of locality asserts that observables in causally disjoint regions are always compatible. Putting the principle of locality and axiom of microcausality together we see that Haag and Kastler, by not providing a detailed physical motivation for microcausality as an axiom, are assuming that spacelike separated regions are causally disjoint. This assumption gets into trouble when you try to sort out precisely what constitutes a causal relationship. We know from the no- signalling discussion that there are some classes of correlation (outcome dependence) that, while they cannot be manipulated, carry many signatures of 'causality' and can reach across spacelike separation.

The requirement of compatibility between spacelike separated regions was the notion that motivated their investigation in the first place. This provides evidence for the claim that Haag's local commutivity is not motivated and instead it is simply asserted. The purpose of the inaugural paper was not to motivate each axiom, but to present reasonable mathematical implementations of physically reasonable assertions and show that the results are interesting. My purpose all along has been to assess whether or not microcausality is a physically reasonable assertion. That microcausality can be found quietly embedded within the fundamental principle that inspired Haag and Kastler's axiomatic project is a strong indication that it might not

[6]This assumption, while not itself trivial, is often made without statement in the literature I have examined. I have avoided tackling it simply because I am intentionally avoiding notions of 'causality'.

be appropriately motivated.

After summarizing the Wightman axioms Haag remarked, that “We regard [the axioms] as well as subsequent modifications and further structural assumptions as working hypotheses rather than as rigid axioms” [Haag, 60, 1992] The axioms, as originally postulated, are to be regarded as *provisional*, and some of the assumptions “... should be replaced by more natural ones as deeper insight is gained” [Haag, 60, 1992]. This provides further evidence that the axiom of microcausality, as presented originally, was not well motivated. The motivation behind the axiom was secondary to the project of getting AQFT off the ground. Now that the consequences of the axiom are better understood a project such as this one is worthwhile to ensure that the axiom still stands. Furthermore, we should not be surprised that a strong a motivation is not presented when the axiom was first postulated.

Haag and Kastler’s motivation for postulating the axiom of microcausality is not sufficient to preserve the axiom beyond the early stages of the project. The motivation lacks mathematical rigour and it assumes that space-like separated regions are necessarily causally disjoint. Furthermore, it ignores other possible axioms that could secure the minimum requirements of compatibility just as well as microcausality and was ultimately provisional. The next section evaluates the relativity of simultaneity and hyperplane dependent interpretations of quantum field theory as motivations by examining the relationship between the axiom of interest and these features of spacetime.

2.5 Relativity of Simultaneity and Hyperplane Dependence

Microcausality suggests a form of causal shielding, and the insistence that this is a relativistic phenomenon is (at least partially) an appeal to the locality conditions apparently entailed by the theory of special relativity. The locality implied by SR can be attributed to the relativity of simultaneity.

In SR, if two regions of space-time (U_1 and U_2) are space-like separated then events that occur in one region (A_1) are outside of the forward and backward light cones of events in the other (B_2). In accordance with the relativity of simultaneity there are reference frames from which it appears that A_1 and B_2 occur simultaneously, A_1 occurs before B_2 and in which B_2 occurs prior to A_1.

On the other hand, if the regions in question were time-like separated (as opposed to space-like separated) then events in those regions would be contained within each other's light cones. In this case there is no frame of reference from which these events appear to occur simultaneously. In fact, the ordering of such events are fixed in all reference frames. Anything in the backwards light cone of an event is said to be in its absolute past and anything located in the forward light cone of an event is said to be in its absolute future.

That microcausality applies only to space-like separated regions is also an indication that its motivation might be found in the requirement of hyperplane independence[7]. Unfortunately, as the discussion on superluminal signalling previously revealed, microcausality isn't strong enough to forbid *all* influences from occurring at space-like separation. On the other hand, perhaps microcausality can be described as a minimal requirement for preserving the relativity of simultaneity. After all, the relativity of simultaneity minimally requires the order of space-like separated events to be independent and the standard interpretation of microcausality is just that: the order that the operations are performed in does not matter (to speak operationally).

In the special theory of relativity, the relativity of simultaneity is derived from other properties and assumptions. If the relativity of simultaneity is going to serve as a motivation for an axiom in QFT then we should investigate the postulates which support the relativity of simultaneity. Section 2.1 showed that Lorentz covariance is secured by the Covariance axiom and so we turn to the principle of relativity.

Recall Einstein's explication of the principle of relativity: "If two coor-

[7] A hyperplane of simultaneity is a fixed time slice of space-time. Choosing a hyperplane of simultaneity is equivalent to choosing an inertial frame of reference.

dinate systems are in uniform parallel translational motion relative to each other, the laws according to which the states of a physical system change do not depend on which of the two systems these changes are related to." [Einstein, 1905. From Brown, 74, 2005]. This is often summarized into a prohibition against frame preferencing. The first thing to settle is what exactly the principle of relativity should prohibit privileging. Is it merely restricting the ability to determine if your frame is moving solely by performing experiments within that frame? Does it require that the outcomes of experiments are independent from the frame of reference? Is it the prohibition against observational or theoretical distinguishability of reference frames?

The principle of relativity is, at the very least, more than just a requirement for covariance. Physically, the relativity principle postulated by Einstein, requires any laws of physics that hold in a given reference frame to hold just as well in another frame that is moving uniformly relative to the first [Brown, 74, 2005]. Generalizing this principle, we might say that the fundamental laws of physics are the same in an arbitrary (inertial) frame of reference. This entails that one cannot determine, by an experiment solely contained within a given frame of reference, whether the frame is moving or at rest.

The next question to settle is at what level the principle applies. If the principle is *observationally* oriented then we would be able to prefer frames while still satisfying the principle, as long as that preferencing was not observationally detectable. On the other hand, if the principle of relativity must be upheld *theoretically* then we can never prefer a frame of reference over another, even if is impossible to detect or determine which frame is being privileged in a particular case.

Using an example from ordinary quantum mechanics it is straightforward to illustrate the subtle yet important distinction between the theoretical and observational levels. The collapse postulate of ordinary quantum mechanics is responsible for many counter-intuitive results and interpretive challenges. One common response to these results is to reject the collapse postulate. Interpretations of OQM that do this are collectively referred to as no-collapse interpretations. Two popular no-collapse theories are Bohmian Mechanics and Many Worlds. Both interpretations reject the collapse postulate and treat the wave function as a carrier of ontological content, but each treats the formalism of OQM at a different level.

Bohmian mechanics requires quantum predictions to be upheld at an observational level. The statistics of experiments and measurement outcomes are required to correlate with those predicted by the formalism. However, beneath the stochastic shell is a deterministic core in which a pilot wave deterministically guides the quantum particles along their apparently stochastic trajectories. Observationally, quantum mechanics is correct. At a more fundamental theoretical level something more subtle and distinctly less 'quantum' is going on. According to Bohmian Mechanics particles have definite positions at all times.

The many worlds interpretation also rejects the collapse postulate and privileges the wave function but does so in a way that requires the quantum formalism to be theoretically correct. For believers of this interpretation the world splits into many distinct worlds whenever a quantum experiment occurs such that each possible outcome obtains. On this theory the universe fundamentally behaves in a way that is accurately described by the quantum mechanical formalism. On the Many Worlds account particles only have definite properties (such as position) immediately after measurement, which is precisely what the formalism asserts and no more.

Both Bohmian Mechanics and Many Worlds are consistent with the results of quantum mechanics, and both are fundamentally deterministic no-collapse theories that treat the wave function as having ontological value. Bohmian Mechanics requires the outcome probabilities encoded in the wave function to be upheld at an observational level and no deeper. Theoretically, in the Bohmian picture, if you could determine the pilot wave and the exact positions of the particles, you would know with certainty the outcome of an experiment. Many Worlds requires the outcome probabilities to be upheld at a theoretical level, which necessitates the postulation of infinitely many worlds such that all possible outcomes obtain.

The theoretical level includes the observational level. If a principle must be upheld theoretically then it is also, by necessity, observationally satisfied.

The inverse is not true. The observational level permits a principle to be violated as long as that violation is not 'observable'. This is the case on the Bohmian Mechanics interpretation. Observationally, the quantum formalism is correct. Theoretically, however, particles have definite positions. Whether or not the principle of relativity is an observational or theoretical demand determines how strict our theory must adhere to the principle, which in turn affects the implications of the relativity of simultaneity.

Whether or not the principle of relativity requires experimental (or event) outcomes to be independent of observational frame may at first, to our Newtonian intuitions, seem obvious. After all, if event outcomes are not frame independent then it would seem to be possible that a particular electron could be in two distinct, incompatible, states (up spin and down spin, for example) depending on which frame of reference you happened to observe a measurement on the electron with respect to. This is simplifying the matter too much. Furthermore, if the success of quantum mechanics has taught us anything it should be that our intuitions formed from experience break down at the quantum level.

As an alternative approach to sorting out the connections between microcausality and the principle of relativity, consider the rejection of hyperplane independence. One such theory utilizes the Newton-Wigner localization scheme and is advocated for by Gordon Fleming. The state of a quantum particle, on this interpretation, depends on the hyperplane that the particle is examined from. It only makes sense to discuss properties of systems when the hyperplane under consideration is also explicitly stated. For example, 'what is the polarization state of this photon?' is an incomplete question. Instead you need to ask, 'what is the polarization state of this photon on this particular hyperplane?' [Maudlin, 208-212, 1994].

Localization refers to the way in which we define objects or systems to be contained within a finite region. The goal of a localization scheme is to appropriately tie operations to regions of spacetime. The standard way to localize operations is via the Reeh-Schlieder theorem. Roughly, the standard localization scheme defines a state ψ to be localized to a particular region R if ψ is measurable within R. Given this definition of localization the

Reeh-Schlieder theorem shows that for any net of local observables satisfying the axioms of Isotony, Covariance and Weak Additivity, the vacuum state is cyclic [Halvorson, 2000]. The cyclicity of the vacuum state has a few counter-intuitive implications; most notably this result allows any state on a Minkowski spacetime to be be approximated by measurements on its vacuum state [Ruetsche, 2010]. Thus, one can approximate the entire vector space of space-time by performing a single operation on a localized region of the vacuum [Halvorson, 2000].

The argument against the Reeh-Schlieder localization scheme leverages this (and other) counter-intuitive results that follow from the theorem to conclude that it violates the 'spirit of relativistic causality'. Although Halvorson notes that "... once one makes the crucial distinction between selective and nonselective local operations, local cyclicity does not obviously conflict with relativistic causality" [Halvorson, 2000], proponents of alternate localization schemes continue to accuse the Reeh-Schlieder theorem of violating the spirit of causality.

On the other hand we have the Newton-Wigner localization scheme which was developed in part to side-step the (undesirable) consequences of the Reeh-Schlieder theorem. To paraphrase Fleming and Butterfield, one could define Newton-Wigner localization as follows:

> If ψ is a state that is localized in region R and ψ_1 is another state that is localized in region R_1, which is disjoint from region R, then the vectors which correspond to these states *in a given reference frame* must be orthogonal [Fleming and Butterfield, 1999, 113].

Fleming's hyperplane dependent account of quantum field theory is built upon this definition of localization.

The Newton-Wigner representation is, at least partially, motivated by a desire to avoid the consequences of the Reeh-Schlieder theorem and still tie operations to localized regions in an appropriate manner. The approach used by the Newton-Wigner localization scheme is the same localization scheme found in ordinary quantum mechanics [Halvorson, 14, 2000].

One reason for Halvorson's insistence that the Newton-Wigner representation fails to side step the odd implications of Reeh-Schlieder is its failure to satisfy microcausality. He then shows that, *because the Newton-Wigner localization does not satisfy microcausality*, it equivocates on distinct notions of localization. Subsequently, this leaves us without an adequate or sensible way to interpret what it means to localize a region or operation [Halvorson, 2000].

Here we may have stumbled upon a motivation for microcausality. By Halvorson's argument it seems that the axiom of microcausality plays a crucial role in distinguishing notions of localization. On Halvorson's interpretation, Fleming wants to discuss localized properties without leaning on notions of 'measurable within a specific region'. The problem then becomes a matter of expressing how a property contained within a region can be distinct from what is measurable within a region. Microcausality motivates Halvorson's concern. Operationally, it deals with localization (in the sense that space-like separated regions are localized) via the notion of 'measurable within' and without microcausality the Newton-Wigner representation risks permitting causal anomalies [Halvorson, 2000].

Halvorson shows that one can construct specific examples of systems which are localized in the Newton-Wigner sense but violate the axiom of microcausality. Since Newton-Wigner localized regions do not require the observables localized in space-like separated regions to commute, it is impossible to interpret the localized states as 'measurable within' the regions that they are local to. If one did try to interpret the states as measurable within the regions then the failure of microcausality would permit causal anomalies. In the operationalist language, this would mean that the statistics of experiments in one localized region would affect the outcome of measurements in a disjoint (space-like separated) region. Fleming, Halvorson notes, is not without recourse at this stage since his project focuses on the notion of localized *properties* and not states [Fleming, 2004].

How can a property be localized if *it* is not measurable (within) the region it is local to. Halvorson attempts to tease apart the notions of localized properties and measurability within a region to no avail. He does so via an analogy to classical mechanics. He notes two distinct localization concepts encountered there: fixedly-localized (or f-localized) quantities occur when a physical quantity is attached to a point in space (such as a magnetic field of given strength) and variable-localized (or v-localized) quantities occur when a quantity takes "... vectors in physical space as their values" (such as the centre of mass of a spatially extended body) [Halvorson, 23, 2000].

It is easy to imagine v-localized quantities that are both localized to a region and yet not measurable within that region. The centre of mass of a large body is an example of such a quantity. However, since the assignment of v-localized quantities is permanent to a particular region it does not grant us any leeway in thinking about Newton-Wigner localized states, which are not permanently attached to their region of localization [Halvorson, 23, 2000].

There are no examples of f-localized quantities which are both localized in a region and not measurable within that region. Halvorson shows this by example, but the argument can be made generally. As soon as you demand that a quantity is not measurable in a region it ceases to be f-localized in that region. This is because f-localization requires a quantity to be attached to a particular point in space. A physical quantity that is f-localized in a particular region is, by definition, measurable within that region. Since f-localized quantities are the classical analog of Newton-Wigner localized 'properties' and it is not clear what it would mean for a quantity to be f-localized in R but not measurable within R, we must conclude that it is not clear how we should interpret the localization map provided by the Newton-Wigner scheme [Halvorson, 23, 2000]. This result only followed because the Newton-Wigner localization scheme does not satisfy microcausality.

Halvorson's argument is not without fault. It is an operationalist mistake to assume that for something to be a property it must be measurable. Operationally such a claim is obvious; a property is so because it is captured by a particular measurement process. This is not a physical requirement, merely a theoretical (and indeed, interpretive) one. Nothing in physics (or the universe) blocks the possibility of there existing non-measurable properties[8].

Fleming's reply is to reject microcausality using the same kind of argument that Halvorson invokes against the hyperplane dependent interpretation, by rendering the definition 'measurable within' devoid of reference [Fleming, 32, 2004]. When discussing Halvorson's position Fleming notes that "[the principle of microcausality] is, in turn, justified via the insistence ... that the concept of 'being localized within a region Δ' receives its meaning from the supposed more primitive concept of 'being measurable within a region Δ"' [Fleming, 28, 2004].

Halvorson derives the concept of 'localizable within' from that of 'measurable within' when defining universal microcausality and Fleming rejects this inference. By utilizing Halvorson (and Clifton's) own claim that "... no object is strictly localized in a bounded region of space" [Halvorson and Clifton, 20, 2002], Fleming's argument follows: "... if there are no strictly localizable objects, **then there are no strictly localizable measuring instruments**. Furthermore, **there are no strictly localizable parts of measuring instruments that could be regarded as objects in their own right**" [Fleming, 32, 2004, emphasis in original]. From here it is clear to see that the notion of 'measurable within' cannot provide grounding for the concept of 'localizable within' since the idea of localizable measuring instrument is rendered absurd. Subsequently, the principle of microcausality (as Halvorson presents it) now stands without support as Halvorson's argument in favour of the axiom depends critically on the link between these two concepts.

Fleming's argument relies on the prior stated assumption that, since a measurement apparatus is extended in space outside of the region which it is used to investigate, then just because two regions are space-like separated doesn't entail that the measurement apparatus used to investigate those regions are. This assumption assumes that measuring devices are not already taken into account when ascribing the property of space-like separation to two regions. Furthermore, the regression argument from 'non localizable

[8]Such properties would be of little interest if they did not also have some influence on the measurable properties of the world.

measuring device' to 'nothing is localizable' fails to appreciate that all measurements inevitably do have an end point. Even on a primitive operational interpretation, the observer reading the output from a display is the end point of the regress.

Microcausality is the axiom dividing these two localization schemes. With it, the standard scheme and the Reeh-Schlieder theorem are victorious and without it the Newton-Wigner localization scheme has no meaningful opposition. Halvorson leverages microcausality to reject the hyperplane dependent Newton-Wigner representation and the response by Fleming is to reject microcausality. The key player in this debate is the axiom itself, and on its status hangs the fate of hyperplane dependence. If one is inclined to support hyperplane independence then postulating microcausality would be a favourable position.

The rejection of hyperplane dependence would be an adequate motivational foundation for microcausality if we knew for certain that hyperplane dependence fails to describe the world. As long as it remains possible that there are preferred hyperplanes of simultaneity it also remains possible that microcausality is in error. The bigger question here, that determines if a motivation for microcausality can be found in the relativity of simultaneity, is if QFTs should be hyperplane dependent or independent. Furthermore, if it is to be a hyperplane dependent theory then the sense of dependence needs to be determined. At the very least, it must be determined if the dependence is at the observational or theoretical level. If QFT is hyperplane dependent at the theoretical level then the arguments surrounding the Newton-Wigner localization scheme indicate that there is no room for microcausality in such a QFT. If it is merely observationally dependent then there might yet be room to preserve microcausality, depending on the nature of the hyperplane dependence. Either way, since microcausality is the key player in the debate on hyperplane dependence, before utilizing this debate to motivate postulating microcausality we ought to settle the status of hyperplane dependence.

Hyperplane dependence and independence both fail to motivate microcausality. Of course, we might find a foot hold for microcausality in its relation to the primitive implications of the relativity of simultaneity and its usefulness in securing the role of hyperplane independence. In the next section I will investigate a motivation rooted in ordinary quantum mechanics.

2.6 Tensor Product Motivations

Having discussed a variety of motivations founded in the special theory of relativity, let us briefly consider one that is instead inspired by ordinary quantum mechanics. Ruetsche presents this motivation as follows; "... assuming that $\mathcal{O}$ and $\mathcal{O}'$ are different systems each with its own C^*-algebra, we represent their union by a tensor product of those algebras; elements of different components of a tensor product algebra commute" [Ruetsche, 112, 2011]. This is a process borrowed from ordinary quantum mechanics, where distinct systems are represented as tensor products. Halvorson compactly describes how the tensor product structure works in ordinary quantum mechanics:

> "Despite the fact that non-relativistic QM makes no reference to spacetime, it has a footprint of the relativistic prohibition of superluminal signalling. In particular, that state space of two distinct objects is a tensor product $H_1 \otimes H_2$, and their joint algebra of observables is $B(H_1) \otimes B(H_2)$. In this tensor product construction we represent observables for system A as simple tensors $a \otimes \mathbb{I}$ and observables of system B as $\mathbb{I} \otimes b$. Thus, we have a version of microcausality." [Halvorson, 25, 2006]

It seems odd to prescribe any sort of relativistic 'footprint' to quantum mechanics. The role of the tensor product structure in quantum mechanics is as a way to separate distinct systems. Classifying this as a version of microcausality makes the assumption that space-like separated systems are distinct, an assumption challenged in ordinary quantum mechanics by the violation of Bell's inequality, which, as we know from section 2.3, permits correlations at space-like separation. If systems at space-like separation are distinct in some sense, that sense needs to be made explicit for this assumption to be grounded at all.

Recall that kinematic independence requires any two observables represented by C*-algebras to commute [Summers, 202, 1990]. After defining kinematic independence Summers draws a connection between commutation and coexistence, establishing that when two observables commute they can be said to coexist (in the rough and ready operationalist sense of simultaneously measurable) [Summers, 203-204][9]. The commutation properties that follow from the use of the tensor product structure entail a kind of kinematic independence between distinct systems[10].

Returning to Ruetsche's proposed motivation, reason must be provided for utilizing the tensor product structure (instead of something else) to describe distinct systems. The answer very well may be 'no superluminal causation' (or no 'action-at-a-distance'), which leaves the position subject to the same fate as superluminal signalling, which was left wanting in section 2.3. The problems with this motivation arise before one even reaches the arguments roots in special relativistic features. A defender of this position will have to contend with more immediate and more pressing issues of circularity and motivational strength.

In explicating the tensor product position the definition of 'distinct system' is vital to the success (or failure) of the project. If spacelike separated systems are equivocated with distinct systems the position runs afoul of circularity. The position would roughly read like this:

Since we use the tensor product structure to describe distinct systems and spacelike separated systems are distinct then we would use it to describe

[9]The connection between coexistence and commutation only applies to observables which can be represented by projection operators. This restriction has been dropped here for simplicity.

[10] Note that Summer shows that kinematic independence is logically independent from statistical independence, so any conclusions drawn in section 2.2 are not applicable to the tensor product motivation [Summers, Section II, 1990].

spacelike separated systems. Systems described by the tensor product structure have commuting observables. But why do we use the tensor product structure for distinct systems? Because distinct systems are non-interacting and their observables ought to commute. Which is a circular position.

Microcausality is motivated by the use of tensor product structure to describe spacelike separated systems, which is a mathematical entailment and so we need a physical reason for applying that structure. The use of the tensor product structure is motivated by physical reasons that can be traced to an argument in favour of or an explication of microcausality. There may be other physical motivations for using the tensor product structure to describe distinct systems.

Even if we ignore the possible circularity of the argument, the tensor product structure cannot be used to motivate postulating microcausality. The tensor product structure implies stronger independence properties than mere mutual commutation. Microcausality is **just** mutual commutation between space-like separated systems, but a tensor product structure also entails the following result: for every $\psi \in A$ and $\mu \in B$ there exists a $\tau \in A \otimes B$ such that $\tau\ |_A = \psi$ and $\tau\ |_B = \mu$ [Halvorson, 25-26, 2006]. Microcausality cannot be motivated by the use of the tensor product structure because the tensor product is a *stronger* property.

If microcausality was to be motivated by the use of tensor product structure then we should postulate the use of the tensor product structure for distinct systems *instead* of postulating microcausality. If space-like separated systems are also distinct then such a postulate would secure everything microcausality is intended to secure and more. Of course, this would require a clear definition of 'distinct system' and assumes that we desire a stronger axiomatic framework over minimalistic axioms (which is also not certain)[11]. Either way, the use of tensor products to describe distinct systems cannot

[11] Additionally, it is possible that the weaker axiom of microcausality is independent of the other axioms while the stronger tensor product structure is not. Since I am not considering the axioms as a set, I will not endeavour to settle whether the tensor product structure could (or should) replace microcausality as an axiom of QFT.

serve as grounds for postulating microcausality.

2.7 The Fate of Microcausality

An axiom for a physical theory cannot be arbitrary. It must be motivated and its motivation must be, at least partially, rooted in physical arguments. Although some of the motivation for pursuing algebraic quantum field theory is the mathematical rigour, an axiom cannot be purely mathematical and cannot be motivated by its tendency to make our lives (either mathematically or theoretically) easier. The axiom, and its implications, must be true[12] and the physical arguments deployed to motivate the axiom must do so appropriately.

With regards to the axiom of microcausality I have broken down the relevant requirements of the special theory relativity and considered each in turn as a motivation for postulating the axiom. Lorentz covariance is already covered by the Covariance axiom. Independence properties, which microcausality can be cast as one of, are nice properties but are not appropriately related to microcausality to justify its postulation. While postulating microcausality fosters a mathematically closer relationship between independence properties, these relationships are not a physical requirement. No- signalling theorems that require microcausality are not comprehensive enough to permit 'prohibition against superluminal signalling' from motivating the axiom. Likewise for separability which is a weakening of standard no- signalling results and so falls to the same line of argument. Haag and Kastler's original motivation lacked rigour, and rightly so since Haag viewed the axioms as provisional. The use of hyperplane independence to motivate microcausality hinges on the outcome of the debate about whether or not hyperplane dependence is a feature of the world. Finally, the tensor product

[12] In the case of QFT, approximately true. How true an axiom for a physical theory known to be inadequate, such as QFT, can be is an interesting question but not one to be addressed here. QFT is inadequate because it utilizes the special theory of relativity which is made obsolete by the general theory of relativity.

structure is a mathematically stronger property than microcausality and so cannot be leveraged for motivation.

While each possible avenue of motivation failed on its own, for its own reasons, to provide adequate grounds for postulating an axiom, they could be joined together to form a stronger motivational story. From each considered motivation we can find a small reason to postulate microcausality. From the above investigations we know that microcausality implies nice relationships between statistical independence properties, is utilized in theorems that prohibit the most worrying cases of superluminal signalling, secures the independence of hyperplanes of simultaneity and captures the basic demands of the relativity of simultaneity. However, the problems with these motivations individually are not washed out in their union.

The compound motivation only gets off the ground if the entailments of microcausality are actually true. This motivation assumes that statistical independence properties are related in the way microcausality causes them to be and that hyperplanes ought to be independent in the sense that microcausality permits. The debate surrounding the Newton-Wigner localization scheme illustrates that there is still room in QFT for hyperplane dependence, casting a shadow of doubt over the compound motivation.

This review has not considered the axioms of quantum field theory as a complete set, instead focussing on a single axiom in a vacuum. I have shown that the axiom of microcausality is not well motivated on its own. What is clear is that supporters of microcausality must do more than gesture towards no- signalling results or the use of the tensor product structure as motivation for postulating the axiom and dig deeper to find a more stable, rigorous and justified foundation to stand the axiom on. There is still room for a good motivation to be found by considering it as one member of the total set of axioms.

Bibliography

[1] Albert, David. 1994. *Quantum Mechanics and Experience*. Harvard University Press.

[2] Brown, Harvey. 2005. *Physical Relativity: Space-time Structure from a Dynamical Perspective*. Oxford University Press, Oxford.

[3] Chang, Hasok. 2009. "Operationalism". *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/operationalism/#CriOpe

[4] Dimcock, J. "Algebras of local observables on a manifold". *Comm. Math. Phys.* Volume 77, Number 3, 219-228.

[5] DiSalle, Robert. 2009. "Space and Time: Inertial Frames". *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/spacetime-iframes/#ExtRelPri

[6] Fleming, Gordon. 2003. *Observations on Hyperplanes: I. State Reduction and Unitary Evolution.* Preprint. http://philsci-archive.pitt.edu/1533/

[7] Fleming, Gordon. 2004. *Observations on Hyperplanes: II. Dynamical Variables and Localization Observables*. Preprint. http://philsci-archive.pitt.edu/2085/

[8] Fleming, Gordon and Butterfield, Jeremy. 1999. "Strange Positions". In *From Physics to Philosophy*. Ed. Butterfield and Pagonis, University Press, Cambridge.

[9] Fraser, Doreen. 2011. "How to take particle physics seriously: A further defence of axiomatic quantum field theory". *Studies in History and Philosophy of Modern Physics,* 42, 126-135. DOI: 10.1016/j.shpsb.2011.02.002

[10] Haag, Rudolf. 1992. *Local Quantum Physics: Fields, Particles, Algebras.* Springer-Verlag, Germany.

[11] Haag, Rudolf and Kastler, Daniel. 1964. "An Algebraic Approach to Quantum Field Theory". *Journal of Mathematical Physics*, Vol. 5, No 7, 848-861.

[12] Halvorson, Hans. 2006. "Algebraic Quantum Field Theory". http://arxiv.org/abs/math-ph/0602036v1

[13] Halvorson, Hans. 2000. "Reeh-Schlieder defeats Newton-Wigner: On alternative localization schemes in relativistic quantum field theory". *Phil. Sci.* 68, 111-133

[14] Halvorson, Hans and Clifton, Rob. 2002. "No place for Particles in relativistic quantum theory". *Phil. Sci.* 69, 1-28

[15] Maudlin, Tim. 1994. *Quantum Non-Locality and Relativity.* Blackwell Publishing.

[16] Rédei, Miklós. 2010. "Operational Independence and Operational Separability in Algebraic Quantum Mechanics". *Found Phys* 40: 1439-1449. DOI: 10.1007/s10701-010-9447-x

[17] Rédei, Miklós and Valente, Giovanni. 2010. "How local are local operations in local quantum field theory?" *Studies in History and Philosphy of Modern Physics*, 41, 346-353.

[18] Ruetsche, Laura. 2011. *Interpreting Quantum Theories.* Oxford University Press, New York.

[19] Summers, Stephen. 1990. "On the Independence of Local Algebras in Quantum Field Theory". *Rev. Math. Phys.* Vol. 2, No. 2, 201-247.

[20] Wallace, David. 2011. “Taking Particle Physics Seriously: A Critique of the Algebraic Approach to quantum field theory”. *Studies in History and Philosophy of Modern Physics*, 42, 116-125. DOI: 10.1016/j.shpsb.2010.12.001