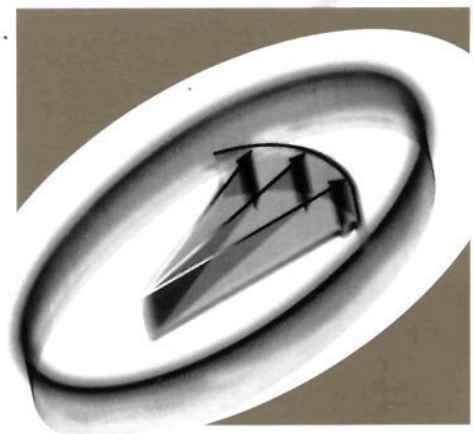


Ontological Aspects of Quantum Field Theory

edited by

Meinard Kuhlmann

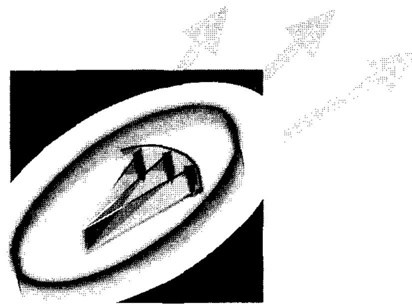
Holger Lyre

Andrew Wayne

World Scientific

Ontological Aspects of Quantum Field Theory

This page is intentionally left blank



Ontological Aspects of Quantum Field Theory

edited by

Meinard Kuhlmann

Department of Philosophy, University of Bremen

Holger Lyre

Philosophy Department, University of Bonn

Andrew Wayne

Department of Philosophy, Concordia University

World Scientific

New Jersey • London • Singapore • Hong Kong

Published by

World Scientific Publishing Co. Pte. Ltd.

P O Box 128, Farrer Road, Singapore 912805

USA office: Suite 1B, 1060 Main Street, River Edge, NJ 07661

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

ONTOLOGICAL ASPECTS OF QUANTUM FIELD THEORY

ISBN 981-238-182-1

This book is printed on acid-free paper.

Printed in Singapore by Mainland Press

Preface

Quantum field theory (QFT) supplies the framework for many fundamental theories in modern physics. Over the last few years there has been increasing interest in ontological aspects of these theories, that is, in investigating what the subatomic world might be like if QFTs were true. Is this a world populated by localizable particles? Or is it best understood as composed of interacting processes? How should we interpret certain non-gauge-invariant structures in QFTs such as ghost fields and antifields? Essays in this volume investigate ontological aspects of QFTs using a variety of physical and philosophical concepts and methods. Although the focus is ontology, these investigations touch on a broad range of topics, from the most general questions of metaphysics to the most minute conceptual puzzles of particular QFTs.

Many of these essays were first presented as papers at the conference "Ontological Aspects of Quantum Field Theory" held at the Zentrum für interdisziplinäre Forschung (Center for Interdisciplinary Research, ZiF), Bielefeld, Germany, in October 1999. This conference brought together physicists, philosophers of physics and philosophers interested in ontology. By all lights the conference was immensely fruitful, in large part because it was successful in promoting a dialogue among participants from diverse disciplinary background. The aim of the present volume is to give contributors an opportunity to advance the dialogue begun at the conference, while at the same time reaching a wider audience. For this reason papers are, as far as possible, written in a way accessible to physicists, philosophers and a general science readership.

We thank the University of Chicago Press for permission to reprint H.

Halvorson and R. Clifton's contribution to this volume. We thank Craig Smith for his assistance in the preparation of the manuscript. We also express our deepest gratitude to Daniel Cartin of World Scientific for his help (and patience!) with the publication of this volume. Many thanks, in addition, to the main organizers of our conference, Manfred Stöckler (Bremen) and Andreas Bartels (Bonn), and also to the people at the ZiF for their kind hospitality. We finally would like to acknowledge the financial support of the ZiF and the Fritz Thyssen Stiftung (Cologne).

May 2002

Meinard Kuhlmann
Holger Lyre
Andrew Wayne

Contents

PART 2: Field Ontologies for QFT

PART 3: Relativity, Measurement and Renormalization

Chapter 1
Introduction

Meinard Kuhlmann
University of Bremen

Holger Lyre
University of Bonn

Andrew Wayne
Concordia University, Montreal

This introduction aims to provide the motivation and background for investigating ontological aspects of QFT, as well as an overview of the contributions to this volume. Section 1.1 explains why one should study ontological aspects of QFT and shows that this investigation will be successful only if concepts and methods of both physics and philosophy are brought to bear. Section 1.2 provides a layperson's introduction to QFT for those unfamiliar with the theory, while section 1.3, aimed at those with some background in physics, surveys the three main mathematical formulations of QFT: canonical QFT in a Fock space, algebraic QFT and the path integral approach, as well as the basic concepts of interacting QFTs. Section 1.4 provides a layperson's introduction to philosophical concepts and methods. Finally, section 1.5 gives an overview of the contributions to this volume.

1.1 Ontology of QFT: Aims, Methods and Work Ahead

Before entering into the more detailed discussion, it seems helpful to consider what it means to investigate ontological aspects of a physical theory like QFT. It will become clear that philosophy and physics need to cooperate on this issue, since the aims as well as the methods of this investigation differ from the usual ones in physics as well as in philosophy. Finally, some examples will illustrate the actual work to be done by the ontologist of QFT.

1.1.1 *Why Study the Ontology of QFT*

In the "Against Philosophy" chapter of his well-known *Dreams of a Final Theory*, Steven Weinberg is deeply disappointed by philosophy's 'unreasonable ineffectiveness' (Weinberg 1992, p. 169) to help physicists, e. g., in their search for a final theory. The only merit Weinberg can see in a philosopher's work is that it sometimes prevents physicists from the evils of other philosophers. Weinberg's attack on the philosophy of science reflects a popular view of philosophy (perhaps more than it reflects Weinberg's own considered opinion). The enemy here, so to speak, is a set of common misconceptions about the aims and results of both physics and philosophy. The misconception about physics is that it, and it alone, is the only legitimate or productive approach to investigating the fundamental aspects of the natural world. The mistake about philosophy is that its methods and results proceed completely independently of the natural sciences. These mistakes seem sufficient reasons to reconsider the very legitimacy of investigating philosophical questions about sciences in general and ontological aspects of QFT in particular.

The aim of ontology is to get a coherent picture of the most general structures of the world, or of *being qua being*—to use Aristotle's famous programmatic description of ontology. One wants to know which kinds of things there are and how everything is related, whether and how some things are composed of parts and whether there are fundamental entities out of which everything else is composed. The ontologist of QFT is then concerned specifically with the story that QFT tells us about the world—provided that QFT is a true theory.

The aim of ontological investigations about physics is unusual for a physicist because ontologists do not aim to make any empirical predictions—at least not in general. Only in some cases will it be possible to decide between two competing ontological interpretations of a physical theory by testing their empirical consequences. At the same time, the aim of ontological investigations about physics, as pursued primarily in the philosophy of science, is unusual for a regular philosopher, too. In contrast to the view in some other philosophical traditions, the development of the special sciences has an immediate and important impact for the philosopher of science concerned with the ontology of QFT. For philosophers of science, the special sciences constitute the yardstick for their philosophical considerations. This applies to both of the two main branches of philosophy of science, the methodological branch, which examines the methods of science, as well as the branch which examines the contents of scientific theories.

One of the most significant philosophers of science who initiated the study of the contents of modern physics was Hans Reichenbach, with his influential enquiry about the ontological status of spacetime (1928). Philosophical, and particularly ontological, investigations into the contents of QFT were initiated by the seminal work of Michael Redhead (1980 and 1983) and have received additional momentum with the anthology by Harvey Brown and Rom Harré (1988), the monographs by Paul Teller (1995) and Sunny Auyang (1995), and the anthology by Tian Yu Cao (1999). The articles in the present volume continue that work, and we wish to emphasize one new feature. We have tried to foster an exchange between philosophers of physics and physicists on the one side, and "pure" philosophers on the other.

Talking about the aims of ontological considerations about physical theories like QFT, one should not forget to address the question of how physicists can profit. In general, philosophy is not in a position to make proposals for answering questions of sciences which are as highly developed as QFT or the general theory of relativity. Nevertheless, ontological considerations can be helpful as heuristics when a theory is not completed yet (as in the case of QFT, partly due to the as yet unsuccessful incorporation of gravitation). Philosophy's potential to be of heuristic value should become clearer by discussing the tools which philosophy uses—of course not primarily *in order* to serve this heuristic aspect. Therefore, the next section will deal with how ontological issues can be investigated with a combination of physical contents and philosophical tools.

1.1.2 *The Ontologist's Toolbox*

Studying the ontology of QFT requires the collaboration of least three different disciplines. First of all and most naturally, QFT itself is needed, which is the object of the study. Second, it is helpful to use the concepts, methods and categorial schemes from ontology and, in the context of QFT, we believe it is the best to emphasize analytical ontology. Third and finally, the place where the contents of science and the methods of philosophy meet is philosophy of science, which is traditionally concerned with topics on the—sometimes floating—borderline between special sciences and philosophy. While more will be said about the relevant philosophical disciplines in section 1.4, this section is concerned with the question how physics and philosophy are related to each other in ontological matters. It will be shown that the major contribution of philosophy consists in its conceptual tools for addressing ontological questions.

Historically there are two diametrically opposed lines of philosophical tradition which, for very different reasons, both reject the view that ontology is something in which physics and philosophy could and should cooperate. According to one line, ontology is mainly a matter for philosophy. It is contended that the structures of being qua being can in the main be investigated by pure thinking in an a priori fashion that is immune to any specific scientific results. The other line of philosophical tradition defends the very opposite attitude towards ontology. According to this point of view, ontology is *only* a matter for physics and other empirical sciences. Only specific sciences, in particular natural sciences, are in a position to say anything about the basic entities there are and their irreducible characteristics. Logical positivists such as Wittgenstein and Carnap hold the view that all that philosophy should aim at is to analyze the structure of language and of scientific theories. As far as the contents of special sciences are concerned, however, philosophy has nothing to contribute, according to proponents of the early period of analytical philosophy. Traditional metaphysics was seen as containing nothing but pseudo-problems based on misunderstandings about the logical structure of language.

Although this linguistic turn has deeply changed philosophy, the impression became stronger over the time that philosophy had deprived itself of some of its most genuine questions, questions that did not all appear to rest on pseudo problems. Moreover, it became ever clearer that the most fierce attackers of metaphysics were doing metaphysics themselves. Around the middle of the twentieth century analytical philosophers such as Quine and Strawson intitiated a revival of traditional ontological questions in an analytically purified way. Eventually a new field of research was established, called 'analytical ontology.'

There are two reasons why analytical ontology is of a particularly high value for the investigation of ontological aspects of QFT. The first reason is that since it is rooted in the tradition of logical positivism it holds empirical sciences in high esteem, and it is particularly used to treating scientific theories. The second reason is the emphasis on logical structures in the analytical tradition, which is of great advantage for the analysis of highly formalized theories of physics. We think that the conceptual tools of analytical ontology are thus particularly useful for the ontological investigation of physics. Some examples may help illustrate this point. Physics can explain which properties a thing has, how its properties evolve in time and out of which parts it is composed. However, it is not a question for physics what a property is, whether the distinction between things and their properties is sensible, how identity and change are to be analyzed and which kinds of part/whole relations there are. While such issues are often irrelevant when looking at the everyday world or at most issues in science, they do acquire importance when the findings of special sciences do not fit into our common schemes any more. Innocent- looking questions like "Is this the same electron as before?" can become unanswerable unless it is understood what sameness and thinghood consist in, and it is exactly these questions which are traditionally in focus in philosophy.

1.1.3 *An Agenda for the Ontologist of QFT*

If one adopts a Quinean conception of ontology[1] it is essential for ontology (as a philosophical discipline) to look for the "ontological commitments" of the "best science available".[2] Although that sounds quite straightforward, there is still a lot of work ahead for the ontologist. In the case of QFT, e. g., it is not immediately clear what the "best science available" is. There are various different formulations or approaches whose ontological commitments seem to be very diverse. For example, while the quantum field is a central entity in the standard Lagrangian approach to QFT, it doesn't even appear in algebraic QFT (see sections 1.3.1 and 1.3.2). The ontologist thus has to examine what the ontological significance of the different formulations is and whether their importance for ontological investigations differs. It could turn out, for instance, that one formulation has a high

[1]It is common practise in analytical philosophy (less common in continental philosophy) to use the notion ontology in two senses. The first sense of ontology is as a philosophical discipline, the second one as a set of fundamental entities. Usually both senses are used without further notice, since in general there is no risk of confusion.

[2]See, e. g., Quine 1948 as well as numerous further publications.

ontological relevance while another one is mainly a convenient device for quick calculations.

Moreover, the "ontological commitments" cannot in general be read off, even if the "best science available" is given. Even if there were no question which formulation to choose or if one were to restrict one's attention to just one approach, it is still not a simple matter to say what the ontological commitments are. To give an example again, in the standard Lagrangian approach to QFT a plethora of different kinds of theoretical entities appear and it is by no means clear which of them lead to ontological commitments and which are just useful for calculating empirically testable predictions. Another example is Feynman's path integral approach to QFT. From a formal point of view, the paths which are so basic in this approach appear to be its primary ontological commitments. However, few people would be willing to understand them in such a realistic manner. But if paths should not be among the ontological commitments of the path integral approach, what are they? Similarly, one can consider the ontological significance of some entities or methods that appear in QFT, such as the vacuum, virtual particles and Feynman diagrams. To give one last well-known example where the "ontological commitments" cannot be read off a given theory, consider relativity theory. Here we have the long-standing debate about the ontological status of spacetime, with the spacetime substantivalists on the one side and spacetime relationalists on the other. Contrary to many people's initial impression, relativity theory does not immediately speak in favor of one of these two traditional points of view.

One successful way to handle this situation seems to us to combine considerations from physics and philosophy. In case of QFT, this means that one first has to review all the different approaches in order to gather and organize the 'material' for the ontologist. The next step is to find various possible alternatives conceptions for an ontology of QFT (and the different approaches to and formulations of it). Among these, the particle and the field ontology are only the best known ones. In considering particle versus field ontologies the well-known discussion about the wave-particle dualism with respect to standard quantum mechanics finds its continuation in QFT. However, there are various further conceivable ontologies for QFT as well, for instance the event interpretation by Auyang (1995). The last—and certainly not uncontroversial—step is to compare the achievements of all the different proposals. An important part in this last stage is to see whether it is possible to exclude certain alternatives. One example of this procedure can be found Malament's much discussed paper "In defense of dogma: Why there Cannot be a Relativistic Quantum Mechanics of (Localizable) Particles."[3] The main contribution of philosophy to these enquiries is to supply conceptual tools, in order to help find and establish coherent ontological theories. The material as well as the final testing of competing ontologies will always be a matter of physics, however.

We think that neither philosophy alone nor physics alone are in a position to paint a coherent picture of the general structure of the physical world that takes all relevant knowledge available today into account. This claim is not tantamount to saying that physics and philosophy could not be pursued without taking notice of each other. All that is asserted is that getting an up-to-date comprehensive idea of the physical world necessitates a cooperation between physics and philosophy. There are legitimate questions about nature which are usually not in focus in the natural sciences.

1.2 A Layperson's Guide to QFT

In this section we shall give a brief motivation of the very idea of a QFT (almost) entirely on a layperson's level. We shall introduce the key notions of a "field", "field theory" and "quantization" and discuss the two possible conceptual routes to QFT proper.

1.2.1 *The Field Concept*

As a useful pedagogic introduction to the concept of a *field theory* one may consider the classical mechanics of a system of (point) masses elastically interconnected with springs. In one dimension this is an oscillating chain with each mass being a little oscillator (and, thus, providing one degree of freedom). Physicists consider this to be a simple model of a system of connected oscillators. If the number of oscillators increases considerably, it turns out useful—for modeling and calculation purposes—to consider the limit of infinitely many masses, i.e. a continuum of degrees of freedom. We then have a vibrating string. Mathematically, its behaviour is captured in terms of a wave equation. Classical wave phenomena, such as water

[3]See p. 1-10 in Clifton 1996. Also see the contributions by Barrett, Dieks and Rédei in this volume.

waves or sound waves or, more complex, electromagnetic waves, are indeed examples of what physicists put under the more general notion of a (classical) field. By definition, *fields in physics refer to quantities with values associated to spacetime points.*

Historically, the field concept arose from the mechanics of continuous media such as fluid mechanics (most prominently Eulerian hydrodynamics). With the advent of electrodynamics, the notion of a field irrevocably became a fundamental notion in physics, since the described entity could not be pictured in purely mechanical terms any longer. The electromagnetic field is indeed an irreducible entity *sui generis.*

There are different kinds of fields, depending on whether the field quantity at any spacetime point is represented with one or many values. The temperature field in a room, for instance, associates only one value—the temperature—to each point in the room. Technically speaking, this is a scalar field, but more generally we may consider vector fields or, even more general, tensor fields.[4] As we shall see in section 1.3.4, the field-theoretic picture of infinitely many degrees of freedom ultimately proves to be the reason for many of the conceptual problems of QFTs regarding the occurrence of infinities and the need for various renormalization procedures.

1.2.2 *Two Routes to QFT*

So far quantum theory has not entered our field theoretic considerations. The natural route to QFT is obviously to quantize a given classical field theory. This is the route which was also first taken historically: the quantization of the electrodynamic field by Dirac in his seminal 1927 paper. In order to solve the problem in a "canonical" way, Dirac took the Hamilton formalism and imposed appropriate commutation relations between the canonical conjugate variables. This may be seen as a general recipe for the concept of quantization: take the canonical variables, make them operator valued and consider their commutators (or anti-commutators). In more formal terms this procedure is described in section 1.3.1. Starting from the classical mechanics (CM) of a particle the route to QFT looks like the

[4]Quantum theory in particular introduces spinor fields as fundamental representations of the Lorentz group.

following:

$$\begin{array}{ccc} & & QFT \\ & & \uparrow \quad CCRs \\ CM & \underset{n\to\infty}{\longrightarrow} & CFT \end{array}$$

There is first the step from CM to classical field theory (CFT) as laid out in the preceding section. This step involves the transition to an arbitrary number of degrees of freedom "$n \to \infty$". Next the quantization in terms of the canonical commutation relations (CCRs) follows.

It is widely recognized that Dirac (in his 1927) invented the method of so-called "second quantization". There was, however, in those early years of the history of QFT a considerable confusion as to which field should actually be quantized. Dirac himself considered the electromagnetic field and clearly distinguished it from the Schrödinger wavefunction, which, purely formally, may be seen as a field, too. However, the latter field's wave equation—the infamous Schrödinger equation—described a mathematically imaginary quantity—in clear contrast to the real electromagnetic field. But the method of canonical quantization applies both to a real classical and the imaginary Schrödinger field. In this sense, second quantization simply means "field quantization". Confusing about the quantization of the Schrödinger field, however, is the fact that this field already stems from a quantization procedure! Thus, the ordinary canonical transition from CM to quantum mechanics (QM) may be referred to as "first quantization" and our diagram can be expanded to:

$$\begin{array}{ccc} QM & & QFT \\ CCRs \quad \uparrow & & \uparrow \quad CCRs \\ CM & \underset{n\to\infty}{\longrightarrow} & CFT \end{array}$$

Most interestingly it turns out that this diagram can actually be closed. The founding fathers of early QFT[5] soon realized the remarkable fact that the quantum mechanics of arbitrarily many particles is mathematically

[5]Besides Dirac mention must be made of the names of Fermi, Heisenberg, Jordan, Klein, Pauli and Wigner. Cf. for instance Pais (1986), Schweber (1994) and Cao (1997) for overviews on the history of QFTs.

equivalent to the QFT framework.[6] The complete diagram thus results in:

$$\begin{array}{ccccc} & QM & \xrightarrow{n\to\infty} & QFT & \\ CCRs & \uparrow & & \uparrow & CCRs \\ & CM & \xrightarrow[n\to\infty]{} & CFT & \end{array}$$

Hence, there are indeed two "canonical" routes to QFT. Loosely speaking these two routes emphasize two possible ontological aspects of quantum fields—the wave or field aspect on the "right-up-route" $CM \to CFT \to QFT$, as opposed to the particle aspect on the "up-left-route" $CM \to QM \to QFT$.

A final remark should be made concerning the impact of relativity theory on QFT. As is sometimes stated, special relativity requires the existence of a QFT, since the requirement of energy-matter transformations—due to $E = mc^2$—on the elementary particle level can be fulfilled in a Fock space framework only where field quanta are created and annihilated by means of raising and lowering Fock space ladder operators (cf. section 1.3.1). Perhaps this should rather be seen a heuristic argument, since, firstly, relativistic quantum mechanics is a viable theory in and of itself.[7] Secondly, the opposite line of argument, the inference from QFT to relativity theory, is not enforced, as we already indicated above: one may very well quantize the non-relativistic Schrödinger field, thus arriving at a non-relativistic QFT.

However, from a fundamental point of view, there most certainly exist deep-seated links between Lorentz invariance on the one hand and the structure of QFTs on the other hand, especially once interacting fields are taken into account (cf. section 1.3.4). The interconnections between Lorentz invariance, gauge invariance and the possibility of renormalization is still one of the frontier questions in modern QFT research—going beyond the considerations in this volume.

[6] More precisely, the n-particle Schrödinger wave function is the transition coefficient $\Phi^{[n]}(\vec{x}_1, \vec{x}_2, ...\vec{x}_n; t) = \langle \vec{x}_1, \vec{x}_2, ...\vec{x}_n; t | n_1, n_2, ...n_n \rangle$ between the abstract Fock space representation (cf. footnote 8) and the localized n-particle state (in the position representation).

[7] One has to explain the existence of negative energy solutions, though, which inevitably occur in Lorentz-invariant field equations, by a suitable interpretation such as Dirac's infamous hole theory of the negative energy sea.

1.3 Basic Mathematical Formulations of QFT

In the following we introduce the three main approaches to QFT: firstly, canonical QFT in Hilbert space (or Fock space), secondly, algebraic QFT (or AQFT, for short) and, thirdly, path integrals. It goes without saying that we can only draw a sketchy picture here and, hence, readers with more deeper interests are referred to standard textbooks such as Ryder (1985) or Weinberg (1995).

To be sure, the three approaches mentioned are only the most common ones among a whole variety of formulations. QFT in Hilbert space usually serves as a first introduction to the field-theoretic formalism. Most QFT textbooks take this approach as a starting point—and most of our authors use it, too. The idea is to generalize and apply the concept of canonical quantization to field theory as opposed to particle mechanics. In contrast to the Hilbert space formalism, AQFT is based on the assumption that the algebra of observables represents the core physical structure of quantum theory. As an attempt to construct QFT on a more rigorous and solid mathematical basis, AQFT quite naturally also permits the discussion of certain foundational issues and, hence, some of our authors appeal to it (Dieks, Halvorson and Clifton, Rédei). For practising high energy physicists, however, path integrals play perhaps the most important role. This formalism is suited to get quick results and its visualization in terms of Feynman diagrams makes it even more attractive. It is, however, not clear whether the path integral formalism allows for a deep philosophical insight into QFT. Since none of our authors makes particular use of it, we shall give only a rather short presentation below.

1.3.1 *QFT in Hilbert Space*

The general scheme of canonical quantization is to formulate the theory in terms of the Hamiltonian framework and to then impose commutation relations between the canonical variables. For ordinary quantum mechanics, the reader may recall the well-known Heisenberg relations $[\hat{x}_i, \hat{p}_i] = i\delta_{ij}$. As a standard introductory model, we shall consider the so-called Klein-Gordon field, i.e. a scalar, real-valued massive field $\varphi(x)$ (with $x \equiv (\vec{x}, t)$) with Lagrangian $\mathcal{L}_{KG}(x) = \frac{1}{2}\left(\partial^\mu\varphi(x)\partial_\mu\varphi(x) - m^2\varphi^2(x)\right)$. Here $\varphi(x)$ figures as a canonical variable analogous to classical position, the canonical momentum $\pi(x)$, conjugate to $\varphi(x)$, is $\pi(x) = \frac{\partial\mathcal{L}_{KG}(x)}{\partial\dot{\varphi}(x)} = \dot{\varphi}(x)$.

As a consequence of canonical quantization, the canonical variables $\varphi(x)$ and $\pi(x)$ become operator-valued $\varphi(x) \longrightarrow \hat{\varphi}(x)$, $\pi(x) \longrightarrow \hat{\pi}(x)$, leading to equal-time *canonical commutation relations* (CCRs) $[\hat{\varphi}(\vec{x},t),\hat{\varphi}(\vec{x}',t)] = [\hat{\pi}(\vec{x},t),\hat{\pi}(\vec{x}',t)] = 0$ and $[\hat{\pi}(\vec{x},t),\hat{\varphi}(\vec{x}',t)] = -i\delta^3(\vec{x}-\vec{x}')$. In analogy to the harmonic oscillator model of ordinary quantum mechanics, where we may represent position and momentum in terms of so-called ladder operators, the general solution of the Klein-Gordon equation (the Euler-Lagrange equation to $\mathcal{L}_{KG}$) can be written as

$$\hat{\varphi}(\vec{x},t) = \int \frac{d^3k}{\sqrt{(2\pi)^3 2\omega}} \left(\hat{a}(\vec{k})e^{-i(\vec{k}\vec{x}-\omega t)} + \hat{a}^+(\vec{k})e^{i(\vec{k}\vec{x}-\omega t)}\right). \tag{1.1}$$

Here, the Fourier coefficients $\hat{a}^+(\vec{k})$, $\hat{a}(\vec{k})$ are to be considered as creation and annihilation operators of momentum states $|\vec{k}\rangle$ with CCRs

$$\left[\hat{a}(\vec{k}),\hat{a}(\vec{k}')\right] = \left[\hat{a}^+(\vec{k}),\hat{a}^+(\vec{k}')\right] = 0, \qquad \left[\hat{a}(\vec{k}),\hat{a}^+(\vec{k}')\right] = \delta(\vec{k}-\vec{k}'). \tag{1.2}$$

The Hamilton operator then becomes $\hat{H} = \int d^3k\, \omega(\vec{k}) \left(\hat{n}(\vec{k}) + \frac{1}{2}\right)$ with the particle number density operator $\hat{n}(\vec{k}) = \hat{a}^+(\vec{k})\ \hat{a}(\vec{k})$. Thus, in QFT the particle number—in principle—is an observable.

The many-particle Hilbert space, the so-called Fock space, which can be defined from (1.2),[8] leads to symmetric states under particle permutation. Hence, canonical quantization amounts to Bose-Einstein statistics. The canonical scheme may therefore very well be applied to the quantization of bosonic fields such as Klein-Gordon fields and gauge fields (for the latter, certain difficulties arise; see 1.3.4). However, if we want to quantize fermionic Dirac fields, the postulate of positive energy necessarily leads to a modification of the CCRs (1.2), which then have to be replaced by *anti-commutation relations.* This points to a deep connection between spin and statistics, as was first noticed by Wolfgang Pauli.

[8] We are looking for a complete orthonormal system of eigenfunctions $|\vec{k}\rangle$ with $\hat{H}\ |\vec{k}\rangle = \omega(\vec{k})\ \ \hat{n}(\vec{k}) + \frac{1}{2}\ \ |\vec{k}\rangle = E(\vec{k})\ |\vec{k}\rangle$. First of all, we define a vacuum $\hat{a}(\vec{k})\ |0\rangle = 0$. Upon application of a creation operator we obtain a particle state $|\vec{k}\rangle = \hat{a}^+(\vec{k})\ |0\rangle$. In accordance with (1.2), the field states are solely characterized by the number $n(\vec{k})$ of field quanta. Thus, we get the *occupation number representation* $|n_1, n_2, n_3 ...\rangle = \frac{1}{\sqrt{n_1!n_2!n_3!...}}\ (\hat{a}_1^+)^{n_1}(\hat{a}_2^+)^{n_2}(\hat{a}_3^+)^{n_3}\ldots\ |0\rangle$ with $\hat{a}_i^+ \equiv \hat{a}^+(\vec{k}_i)$. These states form a basis in the *Fock space* $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3 \oplus \ldots$, which is a direct sum of n-particle Hilbert spaces $\mathcal{H}_n = \bigotimes_n \mathcal{H}_1$ (where $\mathcal{H}_1$ denotes the one-particle state space).

1.3.2 *Algebraic QFT*

AQFT is the algebraic version of axiomatic QFT,[9] a more general attempt to formulate QFT whose emphasis lies on a mathematically rigorous and precise formulation. The ultimate hope is to get a better grip on, or to avoid, the infinities in QFT (cf. sec. 1.3.4). AQFT was invented by Rudolf Haag and co-workers in the 1960's (cf. Haag 1992). It starts from two ideas—the priority of the algebraic structure of the set of observables and a certain notion of locality.

The first prerequisite can be traced back to Pascual Jordan's algebraic reformulation of quantum mechanics in the early 1930's. Recall that, according to the usual Hilbert space formulation of quantum mechanics, observables are associated with self-adjoint operators on $\mathcal{H}$, whereas states are represented by vectors—more precisely, one-dimensional subspaces—of $\mathcal{H}$. The idea of the algebraic approach is to ignore $\mathcal{H}$ altogether and to start with the operator algebra instead. In general this is a *-algebra $\mathcal{A}$.[10]

It is now possible to identify the *physical states* of AQFT with the linear forms[11] ω over $\mathcal{A}$. It can be shown that each ω defines a Hilbert space $\mathcal{H}_\omega$ and a representation π_ω of $\mathcal{A}$ by linear operators acting on $\mathcal{H}_\omega$. This is the famous *GNS-construction* which allows one to reconstruct the Hilbert space from a *-algebra. Mathematically, therefore, Hilbert space and algebraic approaches are equivalent—from a physical point of view, however, the algebraic approach seems to fit more naturally into the ideal of empirical science, since it assumes observables to be primitive.

Now, what about the second premise, the notion of locality? To be sure, the idea of locality has many faces. One is the concept of a "local"—in the sense of "point-like defined"—field in spacetime (as expressed by formula (1.7) in section 1.3.4). Here the assumption is that we may attribute a field value $\psi(x)$ to any spacetime point x. However, from an operational

[9] Axiomatic QFT was initiated by the work of Arthur Wightman (1956); another version of AQFT is axiomatic S-matrix theory (LSZ theory).

[10] Let $\mathcal{B}(\mathcal{H})$ be the set of all bounded, linear operators on $\mathcal{H}$ and $\mathcal{A}$ be a subset $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ with $A, B \in \mathcal{A}$ and $\alpha, \beta \in \mathbb{C}$, then $\mathcal{A}$ is called an *algebra* if $\alpha A + \beta B \in \mathcal{A}$ and $AB \in \mathcal{A}$. If furthermore with $A \in \mathcal{A}$ also $A^* \in \mathcal{A}$, where A^* is the adjoint operator of A, then $\mathcal{A}$ is a **-algebra*.

[11] More precisely, a linear form ω over $\mathcal{A}$, i.e. a mapping $\omega : \mathcal{A} \to \mathbb{C}$ with $\omega(\alpha A + \beta B) = \alpha\omega(A) + \beta\omega(B)$, is called (i) *real* if $\omega(A^*) = \bar{\omega}(A)$, (ii) *positive* if $\omega(A^*A) \geq 0$, and (iii) *normalized* if $||\omega|| = 1$. A normalized positive linear form, then, is called a *state*.

point of view the naive correspondence $x \longrightarrow \psi(x)$ makes no real sense. Our experimental access is always restricted to some finite spacetime region $\mathcal{O}$. One therefore has to smear out the fields such that $\psi(f) = \int d^4x\psi(x)f(x)$ with some smooth test function f of compact support (i.e. the test function vanishes outside $\mathcal{O}$). The collection of all such smeared fields generates an algebra $\mathcal{A}(\mathcal{O})$ and, thus, in AQFT we have to replace the correspondence $x \longrightarrow \psi(x)$ by

$$\mathcal{O} \longrightarrow \mathcal{A}(\mathcal{O}). \tag{1.3}$$

Under suitable physical conditions, $\mathcal{A}(\mathcal{O})$ may be considered the *C^*-algebra* of all bounded operators associated with $\mathcal{O}$.[12]

Thus, AQFT conforms to a notion of locality in the sense of local operators in some finite spacetime region $\mathcal{O}$. In its totality, spacetime can be covered by a *net of local algebras*

$$\mathcal{A}_{loc} = \bigcup\nolimits_{\mathcal{O}} \mathcal{A}(\mathcal{O}). \tag{1.4}$$

In other words, the total algebra of all observables is the union taken over all bounded regions.[13] The above notion of locality in terms of local operators becomes manifest in the *Einstein causality condition*

$$[A_1, A_2] = 0 \quad \text{if} \quad A_1 \in \mathcal{A}(\mathcal{O}_1),\ A_2 \in \mathcal{A}(\mathcal{O}_2) \quad \text{and} \quad \mathcal{O}_1,\ \mathcal{O}_2 \text{ spacelike.} \tag{1.5}$$

The basic mathematical structure of AQFT has now been laid out. However, the link to experimental physics is still a long way off. Whereas physical systems are represented by their observables (operator algebra), the states of a system are given by the linear functionals over the algebra or the algebra's representations, respectively. Now, as we already mentioned, a characteristic feature of QFT is the infinite number of degrees of freedom. Mathematically, such systems inevitably possess unitarily inequivalent representations. Different representation classes, so-called sectors, therefore have to be connected to each other by superselection rules, which include the different, empirically known charges.

[12]A *-algebra $\mathcal{A}$ over $\mathbb{C}$ which is closed in the topology induced by the operator norm $||A|| = \sup_{\psi \in \mathcal{H}} \frac{||A\psi||}{||\psi||}$ and where the involution * and the norm $||\cdot||$ are related by $||A^*A|| = ||A^2||$ is called a *C^*-algebra*.

[13]The net obeys the isotony condition $\mathcal{A}(\mathcal{O}_2) \supset \mathcal{A}(\mathcal{O}_1)$ if $\mathcal{O}_2 \supset \mathcal{O}_1$.

An important result of AQFT is the Reeh-Schlieder theorem, which, roughly, asserts that any physical state may be created from the vacuum. It is exactly here that the notorious quantum nonlocalities (such as EPR-Bell correlations) re-arise, in spite of AQFT's local spirit. The abstract approach of AQFT allows for a whole series of formal proofs—independent from canonical QFT—such as the CPT theorem, the spin-statistics theorem and the Goldstone theorem. It also has ample applications in statistical mechanics.

1.3.3 *Path Integrals*

Let us now, ever so briefly, review the path integral account of QFT. It is based on Richard Feynman's representation of transition amplitudes as sums over path histories weighted by the classical action (Feynman 1948). Its theoretical background is provided by the principle of least action. Classically, the trajectory of a particle corresponds to, and can be calculated from, postulating the action functional $S = \int_{t_1}^{t_2} L(x, \dot{x}, t)\, dt$ to be extremal. The intuitive—or rather, counterintuitive—picture in quantum theory is that the particle actually "travels" on all possible trajectories "at once". It can be shown that the literal mathematical representation of this picture allows one to rewrite the transition amplitude $K(x_f, t_f; x_i, t_i) = \langle x_f, t_f | x_i, t_i \rangle$ between initial state $|x_i, t_i\rangle$ and final state $|x_f, t_f\rangle$ as

$$K(x_f, t_f; x_i, t_i) = \int_i^f \mathcal{D}[x(t)]\; e^{iS[x(t)]} \tag{1.6}$$

with a suitable measure $\mathcal{D}[x(t)]$ in the function space of all trajectories $x(t)$.

This formalism allows a quite straightforward extension to infinite degrees of freedom—and, hence, QFT—and also provides a convenient method of performing perturbation theory needed for the calculation of interactions.

1.3.4 *QFT and Interactions*

So far we have only considered non-interacting fields, but clearly, in reality there is no such thing as a "free field" (how could we ever observe it?). One therefore has to incorporate interacting fields, but then the field equations become non-linear (i.e. they contain coupling and self-interaction terms with two or more fields involved). In order to solve the resulting equations one has to consult perturbation theory, where solutions are approximated in terms of infinite power series in the coupling constant. However, the higher-order expansion will inevitably include divergent terms that are due to the integration over internal loops (where one has to integrate over the internal, circulating momentum). The ultimate reason for these divergences can—as already mentioned in section 1.2.1—be seen in the field-theoretic idealization of localizability in continuous spacetime, which most generally requires a superposition in terms of a continuous Fourier expansion

$$\psi(x) = \int \frac{d^4k}{(2\pi)^4}\ \tilde{\psi}(k)\ e^{ikx} \tag{1.7}$$

with arbitrarily high four-momenta k and, thus, a (non-denumerable) infinite number of degrees of freedom.

In order to cope with the infinities an ingenious procedure, called renormalization, has been developed. The idea is to consider the various coupling constants and masses arising in the Lagrangian as "bare" infinite parameters that refer to arbitrarily small distances. As opposed to this, physical parameters are measured at finite distances. We may therefore think of them as interaction-modified parameters. In QED, for instance, the electric bare charge is screened by vacuum polarization. The procedure of renormalization thus substitutes the bare parameters with appropriate renormalized finite ones. One then has to prove that the particular interaction theories considered can be made finite at all orders. This has in fact been shown for the Yang-Mills theories of the standard model. We shall say here nothing more about renormalization—Huggett's contribution addresses it.

The special question of which particular interactions appear in standard model physics is addressed in this volume, since this refers to the concept of gauge theories (cf. articles of Redhead and Drieschner, Eynck and Lyre). As is well-known from quantum theory, an overall phase transformation does not change the physics, since the expectation values of all observables are left invariant. Such phase transformations are called global gauge transformations. The same is true for local (i.e. spacetime-dependent) gauge transformations. The local gauge symmetry of a free field theory is most generally represented in terms of a gauge-covariant derivative that already includes the correct form of the coupling. This astonishing fact is usually called the gauge principle (cf. O'Raifeartaigh 1995 for an overview of the history of gauge theories).

Gauge transformations clearly refer to unphysical degrees of freedom, and only gauge invariant quantities can be observable. On the other hand, at least the canonical and path integral approaches consider matter fields and gauge "fields"[14] as primitive, since they arise in the Lagrangian. However, these two constituents are decidedly gauge-dependent! Particular technical problems then arise when quantizing a gauge theory, since the redundant degrees of freedom have to be treated carefully. There is also a deep connection between the concepts of gauging and renormalization, which we cannot address here. Nevertheless, the concept of interacting QFT's as gauge theories has been proven one of the cornerstones of modern field physics.

1.4 A Layperson's Guide to Ontology, Semantics and Epistemology

Section 1.1 emphasized the mutual dependence between ontological investigations and physics for our understanding of the natural world. The theme of the current section is that ontological investigations are equally dependent on other branches of philosophy, particularly epistemology and semantics. This section provides an introduction to the philosophical tools and methods used in ontological investigations of QFT, grouped under the topics of ontology, semantics and epistemology.

1.4.1 *Ontology*

Ontology is concerned with what there is, and it naturally includes a range of distinct kinds of projects. At its most general, ontology is usually called metaphysics, and it aims at a "theory of everything," not in the physicists' sense of a theory of all the fundamental constituents of the natural world, but in the widest sense: a high-level, abstract account of *all* entities—from emotions to political systems to galaxies—and the principles governing them. At a more specific level, ontology aims to analyze and systematize our understanding of the entities and processes at work within a specific natural domain. It is at this level that the fruitful interaction between philosophy and natural science can occur (see section 1.1).

[14]In high energy physics, the matter field wave function ψ as well as the gauge potential A_μ (as occurring in the QED Lagrangian, for instance) are called "fields". At least as far as the latter is concerned, the notion of a field should rather be reserved for the potential's derivative, the gauge field strength $F_{\mu\nu}$.

What ontology has to offer is 2,000 years of experience grappling with these issues. Contemporary discussions still begin with Aristotle (384-322 BCE), whose basic categories of substance and attribute continue to provide the paradigm for current ontological thinking. Substances are the enduring stuff of the universe, the underlying constituents which persist and remain the same over time. Attributes are properties that inhere in substances, the way a table has properties of being brown, having a mass of 12 kg, being rigid, and so on. As Simons and Seibt argue in their contributions, it is difficult to make a place for the traditional notion of substance within quantum theories, and well-nigh impossible within QFT. Whether the basic entities in QFT are taken to be particles, fields, events or something else, the traditional notion of substance simply fails to apply.

This leaves a range of alternative ontological frameworks that may be of use in ontological investigations of QFT, ably described and classified by Simons and Seibt. Simons favors a factored ontology which looks for universal invariants that determine the basic ontological categories, or factors. Max Planck is taken as having proposed a factor ontology for physics, where the speed of light in a vacuum, the charge and rest mass of an electron, and Planck's own constant are among the fundamental factors. Seibt argues for the usefulness of a process ontology, one that liberates ontology not just from the notion of substance, but also from the very idea of a particular, thing-like entity. Instead of positing a subject or entity which is undergoing a process or activity, the process or activity is simply basic.

1.4.2 *Semantics*

Semantics is the branch of philosophy (specifically, philosophy of language) concerned with how linguistic expressions, such as words, sentences, equations and theories, get their meaning. It answers questions like: in virtue of what does the creation operator in the Fock space formulation of QFT have the meaning it has? As such, semantics provides one set of tools for bridging the ontological gap between the formalism of QFT and the subatomic world the formalism is supposed to be about.

The most common, and common-sensical, account of how linguistic expressions get their meaning is referential: words get their meaning because they refer to bits of the world. "Snow" has the meaning it does because it refers to, or stands for, snow (the stuff on the ground). Sentences, equations and theories in turn get their meaning from the words that make them up. It is easy to see that on a referential account of meaning, meanings are the links between linguistic expressions and the ontology to which these expressions are supposed to correspond.

Words make use of two quite different routes to refer to the physical world, however, and Auyang's contribution demonstrates that both are operative in QFT. QFTs in contemporary physics are all a species of gauge field theory, and Auyang's focus is on the meaning of gauge field theories. In part, gauge field theories refer to the world directly, in the way that, for example, the proper name Daphne refers to a particular person. In the context of gauge field theory, the spacetime variable x refers directly to a particular label, or numerical identity, for an indexed operator $\psi(x)$. Similarly, the local symmetry group of a gauge field functions directly as a kind (or sortal) concept that individuates a particular gauge field, such as an electron field, from other kinds of gauge fields. Gauge field theories also refer to the world indirectly, or descriptively, the way "the person chased by Apollo" refers to whomever fits the bill. For example, a field's dispersion relation $\hbar\omega_k$ does not refer directly to field quanta—which have no numerical identity and so cannot be the object of direct reference—but rather refers to whatever field quanta fit the bill, that is, have the appropriate wave vector, spin and so on.

Semantics is also useful when looking at another aspect of QFT qua gauge field theory, one examined in the contributions of Redhead and Drieschner, Eynck and Lyre: the role of surplus structure. Gauge field theories are beset by a problem inherent in any mathematical description of physical phenomena: the mathematical structure of physical theory contains surplus structure that does not correspond directly to physical structure. On a referential account of meaning, this surplus structure should thus be empty formalism, meaningless. In gauge field theories, however, the surplus structure includes gauge potentials and ghost fields that can play a curiously active role in constraining the structure the physical field theories can take. Redhead's central question concerns how apparently surplus mathematical structure can tell us something concrete about the physical world. This question asks both about the ontological status of this surplus structure and about the semantical issue of how the structure gets its meaning.

1.4.3 *Epistemology*

Closely related to the semantics of a sentence, equation or theory is the question of how we can gain knowledge of it. One often wants to know not just what a particular statement means, but whether one should believe the statement is true. Epistemology is the branch of philosophy concerned with the nature of knowledge and, in particular, the conditions under which it is justified or rational to believe a statement is true. On the traditional approach, knowledge is defined as justified true belief. To know that "it is sunny today," for example, (i) it must be true that it is sunny today, (ii) one must believe that it is sunny today, and (iii) one's belief that it is sunny today must be justified or rational.

Not surprisingly, it is condition (iii) that is the most controversial, and it will be worthwhile to examine this condition in a bit more detail. Some statements seem to be justified more or less directly by observation; roughly speaking, one is justified in believing "it is sunny today" simply by glancing out the window and seeing sunshine. Philosophers call statements known in this way *a posteriori*. Some statements, by contrast, seem immune to observational justification (or falsification), and philosophers call these *a priori*. What is the justification for believing statements of arithmetic or geometry such as 2+2=4? One does not appeal to any observation or experience here, and equally one is confident that no particular observation in the future will undermine one's degree of belief in these sorts of claims. QFTs, like all scientific theories, contain a rich mix of *a priori* and *a posteriori* elements. Falkenburg's contribution offers a detailed examination of these elements for the case of QFT and draws some interesting conclusions for how epistemic considerations should have an impact on our ontological commitments.

QFT, including its ontological aspects, is for the most part an *a posteriori* enterprise: what there is in the subatomic domain, and what the correct mathematical structures are to describe it, are matters for experimentation and observation largely to determine. The epistemic challenge that faces any claim about QFT, including an ontological one, stems from the fact that these claims are about unobservable elements of the world such as quantum fields, whereas experimental results are typically given to us as observations of localized particle traces indicated by tracks in a bubble chamber, marks on paper, clicks in a Geiger counter or images on a computer screen. And the inferential leap that must be made here from the observable to the unobservable is fraught with epistemic difficulties. This volume includes contributions that focus on two distinct aspects of this problem.

There is the well-known "measurement problem" in quantum theories that arises because the core theory of quantum mechanics predicts that most measurements will not have definite outcomes but rather result in indefinite (superposition) states for the quantities being measured. This problem is compounded in the context of QFT by results, examined in Barrett's and Halvorson and Clifton's contributions, which assert that in relativistic QFT there can be *no* detectable (measurable) objects of finite size—like localized particles! How then can a pure field ontology, containing only infinitely-extended field states, be reconciled with our measurement records of finite-sized, localized particles? Barrett and Halvorson and Clifton explore various responses to this question.

A second epistemic challenge arises out of the complex role that renormalization techniques play within the theory. One of the aims of QFT is to predict experimental results such as the scattering amplitudes (cross-sections) from particle collisions in high-energy physics experiments, and the techniques of renormalization help physicists to do this. However, as Huggett explains in his contribution, these techniques involve steps that, although formally legitimate, raise doubts about the epistemic status of QFT. Specifically, renormalization seems to involve non-deductive jumps or modifications to QFT. The situation is one in which a phenomenological theory (used to calculate cross-sections) fails to be a deductive consequence of the fundamental theory (QFT). QFT, strictly speaking, doesn't have any experimental consequences and so we can never have knowledge of it (or its ontology).

The moral of this section is that if ontology without physics is one-eyed, ontology without semantics and epistemology is blind. Clearly, any ontological claim about QFT relies on a semantics (including a theory of how the claim gets its meaning) and an epistemology (including an account of the justification or rationality of the claim). In this way the tools and methods of ontology, semantics and epistemology are used together to produce the fruitful work this volume presents on ontological aspects of quantum field theory.

1.5 Overview of this Book

Contributions to this book are grouped thematically into four parts, each part focusing on a particular aspect of the ontology of QFT.

1.5.1 *Approaches to Ontology*

Part I focuses on how the methods and results of analytic metaphysics may be of help in elucidating ontological aspects of QFT. The theme of all the papers in this part is that there is potential for an extremely fruitful exchange of ideas between those working on philosophical topics such as metaphysics, ontology and reference, on the one hand, and those working on QFT on the other.

Peter Simons' contribution begins by placing the inquiry into an ontology for QFT within a larger philosophical perspective reaching back to Aristotle. Simons reviews several ontological theories as possible frameworks for QFT, beginning with the traditional substance/attribute theory and then describing set theory ontologies, fact ontologies, process ontologies, trope ontologies, possible worlds and factored ontologies. Factored ontologies may be the least familiar to the reader, but it is this approach that Simons advocates as the most promising for QFT. Simons concludes that what is needed to determine whether a factor ontology, or indeed any other ontology, is fruitful for QFT is a continued exchange of ideas between those working on analytic metaphysics and those working on QFT.

Johanna Seibt begins with a methodological survey that usefully clarifies what an ontological investigation consists of, what results it should yield and what kind of explanations about the physical world it can underwrite. She describes the dominant paradigm in analytic ontology in terms of a set of "characteristic Aristotelian presuppositions" that define the key notions of substance and attribute. Seibt's main critical focus is to liberate ontology from the "myth of substance," the belief that ontology must begin by positing an Aristotelian substance. She argues that substance-based ontological approaches to QM and QFT are seriously inadequate, while trope- and event-based alternatives founder on a conceptual incoherency. Seibt recommends a radical break from the myth of substance. She develops a version of process ontology, *axiomatic process theory*, and suggests that it is the most promising framework for research into ontological aspects of QFT.

Meinard Kuhlmann's commentary compares and contrasts the approaches of Simons and Seibt. Kuhlmann raises specific worries about both approaches. With respect to Simons' factor ontology, Kuhlmann objects that, with more than 3,000 fundamental combinations, this is an extremely unparsimonious ontology that will fail to provide a basis for explanation and understanding of QFT. As for Seibt's process ontology, Kuhlmann points out that it is not clear what, in QFT, could possibly constitute the processes basic to the ontology. Overall, however, Kuhlmann finds merit in both Seibt's and Simons' contributions to the extent that they help resist the dominance of substance approaches and help sharpen the focus on methodological considerations relevant to constructing an ontology for QFT.

How do field theories refer to the physical world? This is the question motivating Auyang's contribution. She contends that questions about reference, central to the philosophy of language, function as a useful bridge between the theoretical formalism and the ontology of quantum field theories. Auyang's Kantian moral is that even our ordinary discourse about QFT presupposes much more than is apparent at first glance, and analyzing these presuppositions is a useful aid to an ontological analysis of QFT.

1.5.2 *Field Ontologies for QFT*

The three contributions of Part II all take, as their starting point, the interpretation of canonical QFT developed by Paul Teller (1995). Teller's main aim in this book is to refine our understanding of two of the central concepts used in work on ontological aspects of QFT, namely the notions of *particle* and *field.* Teller emphasizes the differences between classical particles and quantum particles, particularly the lack of "primitive thisness" in the latter, and he develops the quantum particle concept at length in the context of a Fock space formulation of QFT. For fields, Teller's target is a naïve view about canonical QFT that takes the spacetime-indexed set of quantum-mechanical operators, the operator valued quantum field (OVQF) $\Phi(\mathbf{x}, t)$, to represent a quantum field, in analogy with how spacetime-indexed scalars, vectors and tensors represent classical fields. One of Teller's central arguments is that this naïve view is a wrong-headed approach to the ontology of QFT, and that the OVQF is better understood as a set of determinables (or variables) rather than values of a physical quantity.

Andrew Wayne argues that Teller's characterization of the OVQF as being made up of determinables is based on an unduly restrictive conception of what can count as the value of a physical quantity. He develops an ontology for canonical QFT in which vacuum expectation values (VEVs) play a central role: VEVs for field operators and products of field operators in models of canonical QFT correspond to field values in physical systems containing quantum fields.

Gordon Fleming comments on Teller's claims about particles and fields. Fleming resists Teller's argument that the lack of primitive thisness in quantum particles necessitates a move from a labeled tensor product Hilbert space formalism (LTPHSF) for QFT to a Fock space formulation. According to Fleming, the excess formal structure of the LTPHSF does not have the ontological (or epistemic) costs Teller associates with it. With respect to Teller's claims about quantum fields, Fleming is broadly in agreement with Wayne: for a number of reasons, the OVQF *does* correspond more closely to a classical field than Teller admits.

Teller's contribution usefully develops his interpretation of the OVQF by articulating two further significant differences between classical and quantum fields. The first difference is with respect to their modality. The determinate field configuration of a classical field is physically contingent, in the sense that an alternative set of field values is possible. Not so for the OVQF, since the field configuration here is physically necessary—no alternatives are physically possible. Second, classical fields act as causal agents, producing and explaining observable phenomena. The OVQF, Teller argues, does not act in this robust sense but ought to be interpreted as having only a structural efficacy; like the classical Newtonian gravitational potential field, the OVQF does no more than specify the structure of physically contingent possibilities.

1.5.3 *Relativity, Measurement and Renormalization*

Measurement requires there be some localized object or trace as the measurement record. The special theory of relativity, however, seems to impose certain constraints on QFTs that preclude the existence of localized particles. And renormalization challenges the very status of QFT as a fundamental theory. These interconnected issues form the foci of the papers in Part III.

Jeffrey Barrett asks not what we can do for an ontology of QFT, but rather what an ontology of QFT can do for us. More specifically, he explores what kind of ontology of QFT might help to solve the measurement problem in quantum mechanics and quantum field theory. The quantum measurement problem arises because the core theory of quantum mechanics predicts that most measurements will not have definite outcomes but rather result in indefinite (superposition) states for the quantities being measured. This problem is compounded in the context of QFT by Malament's theorem, which asserts that in relativistic QFT there can be no detectable (measurable) objects of finite size. A pure field ontology, Barrett suggests, is consistent with this result and so might help solve at least one aspect of the measurement problem, although as he points out our measurement records typically consist of objects of finite size (marks on paper, etc.) and not infinitely-extended field states. Such a field ontology is still subject to the original quantum measurement problem, and Barrett's conclusion is that finding an ontology that helps with this problem is a precondition for developing an adequate ontology for QFT.

Hans Halvorson and Robert Clifton's contribution also focuses on Malament's theorem. They generalize Malament's result and respond to a number of objections that have been made to it. They present two new no-go theorems, one against localizable particles that does not assume the equivalence of all inertial reference frames (as does Malament's theorem), and a second against the possibility of even unsharply localized particles in relativistic QFT. They argue, however, that these results do *not* support a field ontology over a particle ontology for relativistic QFT. It is possible, they contend, to develop a particle ontology for relativistic QFT in which localized particles are supervenient on underlying localized field observables.

Dennis Dieks explores ontological aspects of QFT from within the framework of algebraic QFT (AQFT). The central challenge for AQFT approaches is to render them compatible with an ontology of localized events and objects, since as we have seen, accounting for these localized entities is crucial for the interpretation of QFT. Dieks develops a perspectival version of a modal interpretation of quantum theory in which quantum properties are irreducibly relational: a system's properties are defined only with respect to a reference system. Applying this and a basic decoherence condition to algebraic QFT yields the result that definite physical magnitudes can be associated with each spacetime region—precisely the spacetime localized events that were sought.

Brigitte Falkenburg approaches the issue of measurement and ontology from a Kantian perspective, examining the concrete ways in which measurements are made in QFT experiments and the structural features of the empirical reality we can infer from them. She reviews the kinds of experimental evidence gathered in high-energy physics experiments related to quantum electrodynamics and the standard model, and argues that these commit us to a thoroughly relational ontology for QFT.

All discussion of ontological aspects of QFT presupposes that QFT is at least a candidate for a fundamental theory of the quantum domain. As Nick Huggett explains, renormalization techniques challenge this role since they seem to involve non-deductive jumps or modifications to QFT, thereby breaking the connection between QFT and experimental confirmation. The situation seems to be a paradigm case of what Nancy Cartwright has characterized as "how the laws of physics lie." Cartwright's conclusion is that in cases like this, the fundamental theory isn't about the natural world at all; it simply isn't "true". Huggett argues that the renormalization group gives a powerful framework within which to reconceptualize the renormalization process and thereby restore a deductive link between fundamental QFT and experimental predictions.

1.5.4 *Gauge Symmetries and the Vacuum*

QFTs in contemporary physics are all gauge field theories, and this raises deep interpretive puzzles about the relation between gauge field theory formalism and the physical world (see section 1.4.2). QFTs also include a vacuum state that, in contrast with classical notions of the vacuum, is a state with a rich structure, full of energy and potentialities. Contributions to Part IV examine ontological aspects of the vacuum state and ontological aspects of QFTs *qua* gauge field theories.

Michael Redhead asks: How can apparently surplus mathematical structure in gauge field theory tell us something concrete about the physical world? Redhead canvasses the advantages and disadvantages of three approaches to this question. One can invest the surplus structure, such as the electromagnetic gauge potential, with physical reality (and claim that it was not really surplus at all). One can formulate the theory using solely gauge-invariant quantities, such as the loop space approach (using holonomy integrals) in electromagnetism, thus eliminating the surplus structure entirely. However, the route actually taken in the development of gauge theories is to introduce further non-gauge-invariant surplus structure such as ghost fields and antifields. The difficult interpretive issue of the ontological status of this surplus structure remains, Redhead concludes, the most pressing problem in current philosophy of physics.

In their comment, Michael Drieschner, Tim Oliver Eynck and Holger Lyre defend the second of Redhead's options. They introduce equivalence classes of gauge potentials, which they dub *prepotentials*, into the ontology of gauge field theory as basic entities (this is equivalent to a loop space approach). The main drawback of this approach, for Redhead, was its inherent non-locality. Drieschner, Eynck and Lyre argue that, in fact, all of the three approaches conform to a certain sort of non-separability in the sense familiar—but nevertheless different—from EPR correlations in quantum mechanics.

Simon Saunders explores ontological aspects of the vacuum in QFT through the question of whether the energy it contains, the so-called "zero-point energy", is real. Saunders puts his discussion in historical perspective with an account of the development of our vacuum concepts, specifically the classical ether of the 19^{th} century and Dirac's concept of the vacuum as a negative energy sea of particles. Saunders examines the various roles that the Casimir effect has played within arguments for the reality of the zero-point energy, and he delineates the links between the vacuum zero-point energy and the traditional problem of the cosmological constant.

Miklos Rédei approaches the vacuum from within the framework of algebraic QFT (AQFT). A remarkable feature of AQFT is what Rédei calls it "ontological silence": The axioms of AQFT do not include any concepts of field or particle. Rédei argues that this ontological silence about fields and particles does not imply an ontological neutrality, and in two ways. First, AQFT is compatible with the field concept, and there is a two-way derivability relation between quantum fields and local nets of algebras. However, AQFT is not compatible with the postulation of localizable particles in QFT due to considerations stemming from the Reeh-Schlieder theorem and Malament's theorem. Second, AQFT is a causally rich theory, in the sense that it is compatible with common cause principle that posits a common cause lying in the intersection of the backward light cones of any two spacelike separated events. Whether there actually *are* such common causes in QFT is an important ontological question that requires much further study.

References

S. Y. Auyang (1995): *How is Quantum Field Theory Possible?* New York: Oxford University Press.

H. R. Brown and R. Harré, editors (1988): *Philosophical Foundations of Quantum Field Theory.* Oxford: Clarendon Press.

T. Y. Cao (1997): *Conceptual Developments of 20th Century Field Theories.* Cambridge: Cambridge University Press.

T. Y. Cao, editor (1999): *Conceptual Foundations of Quantum Field Theory.* Cambridge: Cambridge University Press.

R. Carnap (1931): Überwindung der Metaphysik durch logische Analyse der Sprache. *Erkenntnis*, 2:219–241. (Engl. transl.: The elimination of metaphysics through logical analysis of language. In: Ayer, A. J., editor (1959), *Logical Positivsm*, Glencoe: Free Press, London: Ill. Allen & U.)

R. Clifton, editor (1996): *Perspetives on Quantum Reality: Non-Relativistic, Relativistic, and Field-Theoretic.* Dordrecht: Kluwer.

P. A. M. Dirac (1927): The quantum theory of emission and absorption of radiation. *Proceedings of the Royal Society of London*, A 114:243–256.

R. P. Feynman (1948): Space-time approach to non-relativistic mechanics. *Reviews of Modern Physics*, 20:367–385.

R. Haag (1992): *Local Quantum Physics.* Berlin: Springer. (2nd edition 1996).

L. O'Raifeartaigh (1995): *The Dawning of Gauge Theory.* Princeton: Princeton University Press.

A. Pais (1986): *Inward Bound: Of Matter and Forces in the Physical World.* New York: Oxford University Press.

W. V. Quine (1948): On what there is. *Review of Metaphysics*, 2:21-38. Reprinted in: *From a Logical Point of View*, Cambridge, Massachusetts: Harvard University Press (second, revised edition 1961).

M. L. G. Redhead (1980): Some philosophical aspects of particle physics. *Studies in History and Philosophy of Science*, 11: 279–304.

M. L. G. Redhead (1983): Quantum field theory for philosophers. In: Asquith, P. D. and Nickles, T., editors, *PSA 1982: Proceedings of the 1982 Biennial Meeting of the Philosophy of Science Association* Vol. 2, pages 57–99.

M. L. G. Redhead (1988): A philosopher looks at quantum field theory. In: Brown and Harré (1988).

H. Reichenbach (1928): *Philosophie der Raum-Zeit-Lehre.* Berlin: De Gruyter.

L. H. Ryder (1985): *Quantum Field Theory.* Cambridge: Cambridge University Press. (Second edition 1996).

S. Schweber (1994): *QED and the Men Who Made It.* Princeton: Princeton University Press.

P. Teller (1995): *An Interpretive Introduction to Quantum Field Theory.* Princeton: Princeton University Press.

S. Weinberg (1992): *Dreams of a Final Theory,* New York: Pantheon Books.

S. Weinberg (1995-1999): *The Quantum Theory of Fields.* Cambridge: Cambridge University Press. 3 volumes.

A. S. Wightman (1956): Quantum field theories in terms of vacuum expectation values. *Physical Review,* 101:860–866.

This page is intentionally left blank

PART 1

Approaches to Ontology

This page is intentionally left blank

Chapter 2

Candidate General Ontologies for Situating Quantum Field Theory

Peter Simons
University of Leeds

Abstract. *Ontology is traditionally an* a priori *discipline purveying its categories and principles independently of mere facts, but this arrogance of philosophers has led them into latent or patent incompatibility with good science and has landed them with philosophical aporiai such as the mind-body problem and the universals dispute. So while maintaining the abstractness and systematic universality of ontology it pays to craft one's categories with an eye to the best empirical science, while not necessarily trying to read the ontology off that science. I present desiderata for a systematic ontology and give several reasons why one cannot use physical theory alone as the source of one's* a posteriori *ontology.*
With this in mind I survey six ontological theories as possible frameworks for QFT, four briefly, two at greater length. The first is the traditional substance-attribute metaphysic, which is clearly obsolete, and on which I expend little time. The second is its modern logico-linguistic replacement, the ontology of individuals and sets touted as semantic values in logical semantics. This too falls by the wayside for several reasons. A third is the closely related ontology or ontologies of facts, against which I argue on general grounds. A fourth is Whiteheadian process ontology, which is an improvement over the previous three but still leaves several questions unsatisfactorily answered. The most flexible and promising to date is the ontology of tropes and trope bundles, which I have discussed in several places. After expounding this I reject it not because it is false but because it is neither broad nor deep enough. As a final, sixth alternative, I present an ontology of invariant factors inspired in part by Whitehead and in part by remarks of Max Planck, and offer it as a promising future abstract framework within which to situate the physics of QFT.

In philosophical discussion, the merest hint of dogmatic certainty as to finality is an exhibition of folly.

Alfred North Whitehead

many shall run to and fro, and knowledge shall be increased

(*Daniel* 12: 4)

2.1 Methodological Preamble

From its inception, metaphysics has aspired to maximal generality, to cover all entities and to give the principles governing them all, in their most abstract and general aspects. Thus metaphysics needs a place for everything, including things which physics does *not* aspire to cover, such as the objects of mathematics, biology, psychology, linguistics, sociology, history, engineering, art, religion, and so on, as well as the familiar and prima facie unproblematic entities encountered in the commonsense *Lebenswelt*. Aristotle's science of being *qua* being is intended as a genuine theory of everything, but only every thing in so far as it exists or is a thing: the details may be left to the individual discipline in question. From this point of view, while physics is one of its most important fish, metaphysics also has other fish to fry, and its interests, categories and principles cannot be dictated by physics alone, but most encompass all things. I am thus no methodological physicalist, though I would call myself a naturalist, in two relatively weak senses; firstly, that there is no division in principle between metaphysics and other sciences, but that metaphysical positions can be suggested, corroborated and refuted by empirical discoveries; and secondly in endorsing the view of the world as a natural whole unriven by invidious bifurcations into separate realms of being, whether they be Platonic vs. concrete, spiritual vs. corporeal, or others of that ilk.

Despite this high-level, abstract mandate, metaphysics can nevertheless not be carried on in splendid isolation from the advances of science. It is a condition of adequacy of any metaphysics that it be not incompatible with the well-corroborated results of scientific investigation. The chief problem for the relationship between metaphysics and the special sciences is that of effecting a junction between the categories and principles of the one and the general conditions and regional ontologies of the other. The metaphysician proceeds top-down, proferring a general framework within which to situate more particular sciences, with an eye on the special concepts, objects, laws and requirements of the various sciences. In the case of a clash, there is no universal recipe as to who is right or wrong, simply work to be done.

It is expedient to use the useful pair of words *metaphysics* and *ontology* non-synonymously. I take *ontology* in the sense of Husserl's formal ontology, as giving the *formal* categories under which any thing falls and the *formal* principles governing them. This is supplemented by what Husserl calls regional ontology, which, since his division into regions is controversial, I prefer to call simply *systematics*, understanding this in a sense analogous to, but generalising from, its use in biology. There it designates what George Gaylord Simpson calls "the scientific study of the kinds and diversity of organisms and of any and all relationships among them."[1] In metaphysics, systematics is the scientific study of the most general kinds and diversity of *any* entities whatsoever, and of the most general relationships obtaining amongst them. As such, the desired junction between ontology or general metaphysics on the one hand and physics on the other is part of metaphysical systematics.

It is because an ontology crafted specially to cope with the needs of QFT, or better of physics as a whole, cannot be expected to provide categories covering the plethora of things such as Baroque music, the life-cycle of nematode worms, the English novel, the rise of Nazism, the evolution of birds, the content of my dreams, the meaning of this sentence, and Fermat's Last Theorem, that a bespoke ontology tailored solely to provide a framework for physics will fail at least to give the whole picture. I do not take seriously any suggestion that such things could all be described, either actually or in principle, without leaving the vocabulary and principles of physical science. At the rather basic level of life, and perhaps even in chemistry, there is no reduction: perhaps the simplest proof of this is that while the bases of DNA each obey the laws of physics, the juxtaposition of bases in the nucleotides is physically contingent, so the information content of DNA and the way it serves to encode instructions for constructing proteins is not governed merely by the laws of physics.

If a bespoke ontology for QFT is bound to be inadequate outside physics, can the quantum physicist get an ontology off the philosophical peg? Let

[1]Simpson 1961, 7.

us consider several popular philosophical ontologies which already exist, with regard to their fitness for providing a formal framework embracing the objects and principles of quantum field theory.

2.2 Substance and Attribute

For much of its history metaphysics has been dominated by the Aristotelian distinction between and theory of substance and attribute. Though no longer as rigidly Aristotelian as in medieval metaphysics, this distinction in one form or another is still a commonplace among philosophers of many persuasions. We may take substance in many senses but I shall consider just two. Firstly, substances as the enduring continuants which bear and survive change and remain numerically identical through such change, from their inception to their demise. Call this the strong sense. This notion of substance has been under increasing pressure since the advent of modern science: it was severely mauled by Hume, Mach and other empiricists, rescued by Kant and Kantians as at best an indispensable form of thought, and has been severely attacked by modern metaphysicians such as Carnap, Quine, Smart and Lewis as incapable of adequately explaining change. While I do not subscribe to all the attacks, I consider that the Aristotelian notion of substance as a basic ontological category is dead and not to be resurrected.[2] One cardinal reason for the demise of substance is the loss of the notion of identifiable and reidentifiable individuals in quantum theory, noted half a century ago by Schrödinger[3] and highlighted in much writing in the philosophy of physics. There is a much weaker sense of 'substance' which is beloved of logico-linguistically inclined philosophers, according to which substances are simply individuals, that is, entities of the lowest logical type, as distinct from their attributes (properties and relations) and any attributes of higher order. Substances, or individuals, in this sense, need not be continuants, indeed in the four-dimensionalist ontology of those philosophers just mentioned they are not. Nevertheless they are still individual, countable entities of lowest logical type.

While this notion is much more generally applicable because of its greater flexibility, and its more obviously pseudo-linguistic character, it is

[2]Simons 1998.

[3]Schrödinger 1950.

still radically undermined by the fact that in states of physical systems as described by QFT, there is no definite answer to questions about how many particles of a certain kind are to be found in a given spatiotemporal region. Whatever particles are, they are not little individual substances. And if they are attributes, what are they attributes of? What separates them into different instances? Neither space-time nor anything else is attractive in the role of an independent substance.

In retrospect it is clear that confidence about substance in the weak sense depends on a tacit assumption about the relationship between language and the world, a thesis which I call the pre-established harmony of linguistic and ontological categories. It is assumed either without argument, or else under various kinds of non-realist banner, that no wedge can be driven between the general syntactic categories of name, predicate and sentence, and the general ontological categories of thing (substance), attribute and state of affairs. There may not be a one-to-one correlation between expressions and designata, but there is a correlation of categories, syntactic with ontological. I submit that in the light of QFT and other modern science, as well as for general logical and metaphysical reasons, there is no good reason to suppose such a harmony obtains. The harmony thesis is unacceptably anthropocentric.

2.3 Applied Set Theory

This section would be unnecessary were it not for two things. One is that the standard semantics for logical languages is couched in terms of sets, so the view is familiar. The other is that justifiably highly respected ontologists such as Quine hold to a view whereby it is possible that all that exists is sets.

Consider the first, less radical view. Standard logical semantics starts from a domain of individuals, taken as ur-elements, and construes attributes as sets of ordered tuples of these. All other objects needed for a semantics are manufactured out of sets, with the possible exception of truth-values and, in modal semantics, possible worlds. To these I shall return later. Here I want to concentrate on the sets. Set theory is widely used, well-understood, popular, and familiar to philosophers who use or rely on mathematics, and for these reasons it tends to drive better alternatives from view.

Set theory is an invention of mathematicians interested in infinite numbers. It is barely more than a century old, which makes it a veritable infant by comparison with substance theory. The structural richness of the theory ensures that it can encode many different structures isomorphically. But to encode is not to identify. The human race could be encoded as a set of ordered pairs of sequences of numbers, of which all but the first are from the set 1,2,3,4 standing conventionally for the fours bases of DNA and the first encodes position in a litter of twins, triplets etc. Since apart from clones each human's DNA sequence is unique, and the ordered pairs would represent parenthood, this would do the trick. It would however be patently absurd to identify human beings with sets of pairs of sequences of numbers. So it is with set-theoretic semantics: while we may represent (for some, limited purposes) attributes as sets of tuples of individuals, that does not mean they are these: attributes are causally active or at least causally relevant, e.g. for our senses, but sets, being abstract, are causally inert.

I said it was absurd to suppose human beings might be sequences of numbers, and so abstract sets. Yet Quine comes close to such absurdity when he seriously suggests[4] that physical bodies in general may be considered as sets of occupied spacetime points, these as sets of quadruples of real numbers, and real numbers finally as sets constructed within pure (without urelements) set theory. One may be forgiven for thinking that this has thrown out the baby with the bathwater, and then thrown out the bath as well! Quine's mistake is rampant structuralism. Structuralism is I think wonderful in its right place: within mathematics. But the real world is not merely structural: it has qualitative content. With these few remarks I shall leave sets aside as unserious metaphysics, despite their popularity among logicians.

2.4 Fact Ontologies

Ontologies embracing facts or states of affairs are also increasingly popular again after a period in the doldrums. Their principal contemporary advocate is David Armstrong[5] though he was preceded by Roderick Chisholm.[6]

[4]Quine 1976.

[5]Armstrong 1997.

[6]Chisholm 1976.

The heyday of fact ontologies was in the period 1890–1930, when they informed the formal ontologies of the logical atomists Russell and Wittgenstein, such diverse British philosophers as Bradley, McTaggart and Moore, and Austro-German philosophers like Marty, Stumpf, Meinong, Husserl and Reinach. In general they supplement the thing-attribute picture mentioned before, with states of affairs being taken as special complex entities somehow composed (Armstrong says: non-mereologically) of an attribute and the right number of terms. States of affairs are felt to perform jobs other entities cannot, in particular being truthmakers for atomic and perhaps other propositions, perhaps objects of veridical propositional attitudes, and perhaps terms in the causal nexus. I shall be brief with fact ontologies. Though the idea of using entities in a multiplicity of roles is admirable, it seems to me, and I have argued this extensively elsewhere,[7] that every role for which facts are invoked can be played to as great an advantage or more by some other entities to which one would in any case be sensibly committed. Therefore by Ockham's Razor there is no need to postulate facts or states of affairs.

2.5 Occurrent Ontologies

Occurrent ontologies have been generally less commonly accepted among philosophers than substance ontologies but their acceptance has been increasing of late because of arguments by philosophers such as Carnap, Quine, David Lewis and Armstrong to the effect that change cannot be adequately described within a substance ontology. The position is that the metaphysically basic items are four-dimensional occurrents, not three-dimensional continuants. There are several flavours of occurrent ontology, depending on whether continuants are denied altogether, are construed as simply occurrents although wrongly thought of as lacking temporal extension, or are regarded as secondary, derivative entities based on occurrents. I have argued elsewhere[8] and will not repeat here, that only the third option is reasonable: this does not deny continuants altogether, but it denies them the fundamental status they have in a substance ontology.

[7] Simons 1997 (2001).

[8] Simons 2000a, b, c.

Occurrent ontologies emphasize the dynamic side of reality and are thus capable of accommodating and emphasizing the dynamics of QFT as describing, at least in part, real processes. Since continuants are, on the view I uphold, secondary objects derived from occurrents, in fact as invariants under certain kinds of equivalence relation,[9] there is no metaphysical *Angst* attached to the apparent loss of individuality of particles in entanglements. In such cases the equivalence relations which standardly allow us to isolate and individuate and treat as countables individual invariant continuants simply do not hold.

Two further features of occurrents render them more welcoming to quantum field theory than substance ontologies: occurrents are not exclusive in their occupation of space but may overlap or coincide spatially for a time, so may intermingle and interpenetrate as waves do. Also, sudden or abrupt transitions, snap events, are also occurrents, albeit briefer ones. Whereas classical physics emphasized the continuity of processes, in quantum physics some transitions are discontinuous and abrupt.

The most thoroughly worked out process ontology to my knowledge is that of Whitehead. The ontology of *Process and Reality* is certainly crafted with an eye to subsuming physical science, though he did not take the radical changes wrought by quantum theory to the heart of his system. However, Whitehead's particular complex process ontology is beset by several peculiar problems. Firstly it enshrines a Platonism of eternal objects which is difficult to square with naturalism. Secondly there is the inadequacy of Whitehead's treatment of mind: in a nutshell, he is a Leibnizian panpsychist. Thirdly, there is an unacceptable duality between what Whitehead calls genetic division and coordinate division. He simply posits these as distinct, does not investigate their common features, and does not expand on the tenuous link of occupation or enjoyment which connects them. Paradoxically, what Whitehead calls process, the concrescence of actual occasions, is not as such a development in time, but a logical selection with only indirect repercussions for temporal process. So while, as I emphasized elsewhere,[10] Whitehead's architectonic of categories provides a valuable lesson on how to craft a radically revisionary metaphysics, the content with which he fills the framework is not up to the task.

Occurrent ontologies which say that all that exists are occurrents are

[9]Simons 2000a, b, c.

[10]Simons 1998b.

mistaken, because they leave no place for those things which are genuinely invariant, such as the charge of the electron (that universal), or such familiar particulars as our sun, with all its vagueness. But occurrents will have a place within *any* adequate ontology: the question is only what place. Both common sense and physics tell us that the world around us is an interplay of the stable and the transient: that which is stable through transience, either temporarily or more permanently, may be conceived as a continuant which lasts as long as the condition holds which may be expressed by a suitable equivalence relation among more transient events. Some more abstract configurations appear to be stable in themselves and their kind: some of these constitute regularities which we treat as universal constants, others as universal conditions or laws.

So occurrents should find a place in an ontology fit for physics, and continuants follow on their coat-tails as more or less temporary stabilities or invariants. It is not settled however whether occurrents are basic entities or not.

2.6 Trope Ontologies

Tropes are (usually, perhaps not always) individuals which by their nature cannot exist in isolation: they are dependent entities. Tropes are most frequently invoked in modern ontology to provide a nominalistically acceptable account of predication: some tropes form kinds which may be considered to correspond to properties and relations, but the tropes themselves are non-repeatable individuals. For example the particular mass of a body at a time is a trope of it: it is dependent, since there is no mass without something which has this mass. Whether a mass is a basic entity is a moot point however. Some trope theories postulate a substratum, a thin particular, which holds together a collection of tropes in a concrete (self-sufficient) individual. I do not think substrata are necessary: a concrete individual may be considered a nexus or complex of tropes (their older German name, *Moment*, is in many ways more appropriate here.) Such a nexus will typically consist of an inner core of tightly co-dependent tropes constituting the individual's "essence" and a corona of swappable or variable adherent tropes allowing it to vary its intrinsic features while remaining in existence.[11]

[11]Simons 1994.

Tropes are one somewhat disorganized corner of a rich panoply of dependent entities in the world, among which real boundaries (of lower dimension than the things they bound) may be other candidates.[12] Points of space, moments of time and point-instants of space-time might be dependent boundary particulars.[13] Tropes may be multiply dependent or relational, my favourite example being a collision between two bodies. The interchange of a gauge particle between two particles offers us a case of relational dependence: the gauge particle is dependent on both others and the interplay of all three constitutes an event in which the gauge particle is a short-lived invariant, whereas the others are invariants which may outlast the interaction.

Qualitative tropes may be illustrated by the colours and flavours of quarks. Note that at the subatomic level the relatively "concrete" entities are not very complex: in traditional terms, they have rather few properties. Indeed the limiting case of a entity consisting solely of a single qualitative trope (a quale rather than a quantum) is not perhaps to be ruled out *a priori*: it would not then be dependent, but it would be (potentially) qualitatively simple, when isolated; such interactions as it might participate in would add further, relational attributes to it, but non-essentially (enriching it rather as a colourless piece of glass may appear red when reflecting a sunset). Tropes which are quantitatively comparable and may have spatial directedness are instances of such properties as charge, spin and isospin.

Trope theory has several advantages over theories invoking universal attributes. Tropes are particulars, and as such localized rather than repeated. They are causally effective. They can be either continuants or occurrents: there are probably both, though the relatively derived status of continuants suggests they are all tropes or trope-like, based on their underlying occurrents. Finally tropes share with processes mutual penetrability and superposibility. A further aspect of the flexibility of tropes which makes them attractive is that they may fuse and divide, retaining or summing quantitative features without the need to postulate persisting individual substances, thus avoiding the problem of so-called identical particles. Indeed I do not see why tropes need to be always little individuals at all: just as the mass/count distinction applies nominally in the distinction between

[12] Though against this see Simons 1991.

[13] Again I personally doubt there are such things: they are mathematicians' abstractions rather than real.

things and stuff, and verbally in the difference between events and processes, so it might apply adjectivally in the distinction between integrally countable tropes and merely measurable trope-masses.

As will be seen in Section 2.8 below however, a tropes-only ontology is insufficient for several reasons.

2.7 Possible Worlds

Possible worlds are a great favourite among philosophers trying to give a semantic account of modal statements. They also find a somewhat different use in quantum physics in the many-worlds interpretation of Wheeler. Now despite their popularity, and without wishing to hide my deep scepticism about them, I intend to leave them on one side for the purposes of this discussion. The reason is that whether you subscribe to processes or substances or tropes or sets or any suitable mixture of these categories, possible worlds are orthogonal to your commitments. You can have possible worlds with such an ontology or you can try doing without. They come in when one is discussing such issues as possibility, inevitability, contingency and the like. I am not saying they are irrelevant to the discussion, for three reasons. One is that in some areas such as probability, which are germane to quantum theory, possible worlds interpretations do purport to offer a way to make sense of the relevant figures and concepts.[14] Another point is that possible worlds and abstract objects have to some extent complementary roles: an anchorage for laws in abstract entities may preclude the need for a possibilist interpretation, while an explanation of regularities in terms of possible worlds may seem to make a Platonist ontology less needful. But possible worlds introduce an extra layer of complexity into the argument and since I would prefer to manage without them anyway if it can be done, I shall not discuss them further.

2.8 Factored Ontologies

Having come this far you will no doubt be wondering what I do think is the right way to go rather than what is not. I have indicated that I think our ontology will need processes, with continuants as some of the transtemporal

[14] Cf. Van Fraassen's interpretation of quantum mechanics: Van Fraassen 1991.

invariants amongst them, and that tropes promise flexibility. But this is surely only partial and cannot yet aspire to completeness. No doubt one can attempt to see how far one can come with a tropes-only ontology, it being understood that the continuant/occurrent duality is orthogonal to this. [15] However in talking about tropes and processes we have been using ontological vocabulary which characterizes these entities: for example processes, unlike continuants, have temporal as well as spatial parts. Tropes are ontologically dependent, and there are several flavours of dependence. If something is dependent, then surely, as Bolzano argued, there must be at least one thing that is independent, even if it is the whole world.[16] So there cannot be a tropes-only ontology: the world is not a trope, and the world is not nothing. Then there is the question of the status of space-time itself. Is it an entity in its own right with its own attributes, or is it something dependent on the relationships among other entities? We have talked of universals and invariants versus particulars and variable features, of qualitative versus quantitative properties, of superposition. Perhaps not all of this talk will translate into corresponding entities, and here is a consideration as to why not all of it can.

Assuming there are several fundamentally different kinds of entity in the world, is there any ontological ground for their difference or is it brute? Are they different for a reason or just "simply different?" Without having any way at present to make it convincing to you, let me propose that there is some reason why different kinds of things, even most general kinds, are different. Suppose there are two fundamental kinds A and B. If the ground is itself another kind of entity C, absent in one and present in the other, or two entities C and D present in A and B respectively but differing and excluding one another, then clearly we have not reached the most fundamental kinds of entities: Cs, or Cs and Ds, are more fundamental. But this was supposed false. So there can be no such kind C or kinds C and D. But there is a ground of the difference. Whatever this is then, it must be (1) not a kind, i.e. it must be a single thing, and yet (2) its "presence" in As must make them As and the "presence" of something else in Bs must make them Bs, and furthermore the particular ground of As and that of

[15]See Campbell 1990 and Bacon 1995 for attempts at tropes-only ontologies. (In Bacon's case it is tropes + sets but he – wrongly in my view – counts the sets as part of the methodological apparatus rather than part of the ontology.)

[16]Cf. Simons and Ganthaler 1987.

Bs must be primitively opposed. These are conditions remarkably similar to those of Platonic forms, which are individual, yet multiply determining of basic kinds.[17] The idea of there being certain universal invariants which determine all ontological categories is not new, but perhaps the depths to which one must dig in order to arrive at them are not generally recognized. Here let me first mention two convergent views. One is that of Max Planck, who at various times cites universal constants as "the really substantial": for example here he is writing in 1910:

> "If however, one must ask, the concept of a mass point, which previously assumed as fundamental, loses its property of constancy and invariability, then what is the really substantial, what are then the invariable building blocks out of which the physical cosmos is put together? — ... The invariable elements ... are the so-called *universal constants*".[18]

Planck cites as examples the speed of light in a vacuum, the charge and rest mass of an electron, the gravitational constant and "his own" constant as examples. He makes similar statements throughout his career. The point here is not to argue from authority, but to find similar ideas which can help to illuminate a position. The universal constants appear not to be classes of individuals: there are not lots of little instances of c or h found all over the world: they are rather constants of proportionality among multifarious quantities exhibited throughout the cosmos and go as it were to make up the texture of the cosmos. As such they are ubiquitous and hence universal, but they are not universal*s* with many instances, rather they are abstract particulars constraining the *form* of physical happenings. They are also probably finite and indeed few in number.

This brings me to a second convergence. What is substantial (i.e. fundamental) and what is not (i.e. what is dependent, or, as I shall say, a moment) according to a given physical theory is quite sensitive to theoretical changes. Are waves or fields fundamental and particles moments or the other way around? Or are both dependent on some third thing? Is spacetime fundamental and matter-energy a moment or the other way round? Or are both dependent on some third thing? One may be forgiven

[17] This argument is put forward in greater detail in Simons 2002.

[18] Planck 1958, 44.

at times as a non-physicist for being confused. There is not one message coming from physics as to its ontology, but several at once. This is another reason for not wanting to craft one's ontology solely on the basis of this or that current physical theory. But a way to reduce and partly finesse this variation is to look for *formal* invariants among the different theories. True, the fundament/moment distinction is one aspect of ontological form, but it is not the only one, though through the prevalence of substance-ontology it has been accorded perhaps exaggerated prominence. The claim that it is ontological structure rather than particular choices of ontological category as basic or fundamental that provides for continuity and convergence is made by Tian Yu Cao[19] and that is congenial to the position I am working towards here.

To return to the outline proposal, the formal features which I suggest any ontology which can aspire to the title of scientifically adequate must have, whatever they turn out to be, are decidedly *not* categories of entity. To fix terminology, I shall call then *basic factors* and the collection of them and their legal combinations and superpositional ramifications *the basic field* (no connection here with 'field' in the physical sense). Anything (simple or complex) that belongs to the basic field is a *basis*. A simple fundamental entity or *element* is constituted when a suitable combination of basic factors is realized together as a *being*. Other beings are more complex, being either elements with other elements as proper parts, or nexus of involvements of elements with one another, or collections (multiplicities, pluralities) or elements or other beings. The collection of all beings constitutes *the ontic field*. The ontic and basic fields exhaust all *items* (to use a neutral word).

All of this tells us nothing about what the basic factors are or how they combine, which is in general something to be decided not *a priori* but by a complex to-and-fro process[20] of hypothesis, construction and refutation or corroboration, where the data from physical theories afford part of the testing material. But some clues are afforded by some of the terms already used: being simple vs having proper parts for example gives one dimension, being one vs being many gives another, being fundamental vs being dependent another. These are factors applicable beyond physics: if truly formal they are ontologically topic-neutral and belong to what Husserl called *for-*

[19]Cao 1997, Section 12.4.

[20]See the Biblical motto at the top of the paper.

mal ontology. I shall come back to a provisional set of factors later.

Historically, factored ontologies are not novel. Empedocles propounded an ontology of four elements: air, fire, water and earth, characterized according to the two pairs of opposed factors hot/cold and wet/dry. Early Aristotle (*Categories* Chapter 2) likewise had four general kinds of elementary entity, characterized according to whether or not they are said of a subject and whether or not they are in a subject. In each of these cases every consistent combination of factors is legal: I say the space of combinations is *orthogonal*. A more recent and non-orthogonal set of combinations is provided by the Polish phenomenologist Roman Ingarden,[21] who calls his factors *existential moments* and the categories thus defined *modes of being*. An ontology, whether factored or not, and the metaphysics of which it is a part, may be descriptive or revisionary, to use Strawson's term.[22] Without arguing the case here, I consider an adequate metaphysics for the future has to be, at least for a long time to come, revisionary. I imagine this position will meet less resistance among physicists and philosophers of physics than elsewhere. I take Whitehead as the paradigm twentieth century revisionist in metaphysics. Whitehead is often dense and it is easy to suspect (wrongly), when perusing one of his more purple passages, that he sometimes lapsed into Heideggerian nonsense. But revisionary metaphysics need not and should not lose its anchorage in common sense and science: they are its "reality checks". Heidegger and his admirers despise these checks and pay the price. Whitehead however both tried to incorporate recent science into his systematics and to lay an architectonic scheme reducing the chances of error, while admitting the futility of any assumption as to finality.[23] This combination of humility and audacity is what modern metaphysics needs. Incidentally Whitehead too recognized that the ultimate is not a kind of being. In his system it is called *creativity*. The first thinker to postulate an ultimate which is not a kind of entity was however Anaximander of Miletus, who called it *apeiron*, the indefinite.

I began by citing Planck on fundamental constants. Basic factors are only analogous to these, since their writ must run across the whole range of metaphysics, and a formal ontology cannot be tied to particular values of magnitudes, which is a material issue of content. The basic factors I enu-

[21]Cf. Ingarden 1964.

[22]Strawson 1959, 9.

[23]Cf. Simons 1998b.

merate below and the systematics built using them owe much to insights of Mayr, Carnap, Wittgenstein, Ingarden, Planck, Whitehead, Frege, Darwin, Leibniz, Suarez, Ockham, Scotus, Aristotle, Plato and Anaximander, but their selection was due not to historical comparisons but to testing against the task of producing a general ontological framework for the design of a potentially omni-representing knowledge representation system for computers, the system called PACIS, which was the result of well over a decade of cooperation between philosophers and software designers.[24]

The Ontek basic factors, called *modes*, come in eleven families of two or (in one case) three modes apiece, the families being called *modal dimensions* and their combinations of one from each family—all combinations being legal, so the system is orthogonal—being called modal configurations, yield in total $2^{10} \times 3 = 3072$ kinds of element, which is richer than most category schemes offer. Here are the modes:

Scheme: Modal Dimension: Mode 1 – Mode 2 (– Mode 3)

(1) Complementarity: thetic – kenonic
(2) Valence: moietic – plene
(3) Inherence: synthetic – aphairetic
(4) Status: heteronomous – autonomous
(5) Bias: aisthetic – poietic – choate
(6) Station: relative – absolute
(7) Objectification: illic – hæccic
(8) Extension: allocate – unicate
(9) Vergence: furcate – bracteal
(10) Sortance: diaphoric – adiaphoric
(11) Incidence: anent – perseic

These Ontek modes are novel in three respects. Firstly, they are relatively numerous and thus give rise to a large multiplicity of possible combinations. As a result the complexities of the resulting ontology are dubiously surveyable by individuals. Secondly, they arose not because of philosophical reflection but in order to provide a formal framework for software capable of representing anything. Whether they are in fact successful in that role

[24]The philosophers are David Woodruff Smith, a phenomenologist, Peter Woodruff, a logician, and myself: the principal software designers are Chuck Dement and Steve DeWitt of Ontek Corporation. A previous short exposition is Simons 1999. PACIS is an acronym for 'Platform for the Automated Construction of Intelligent Systems'.

has yet to be fully tested, but they have yet to "break" in use. Indeed their adequacy or otherwise is independent of whether the software they inform is ever put to general use. Thirdly, they are the product not of one individual but of a team. There is much more to the ontology and systematics developed for PACIS and it has applications outside computing, since it is a metaphysic. I do not expect the names or the rationales for each of these, which have been the outcome of a long and painful process, to be self-illuminating; I mention them only for completeness to show that much work has been done. The outcome of more than fifteen years of trial and error with reference to several fields of application, including database design, computational structure, engineering, corporate organization, systematic biology and common sense, will not be apparent from the list as such. Metaphysically difficult issues such as dependence, identity, object and attribute, whole and part, one and many, function and argument, positive and negative, being and becoming, are reflected in them. I will only say that this is the list of factors which has emerged from the debate and that we posit them as capable of accommodating other formal concepts and framing material concepts. If it turns out that any aspect of secure physics is not representable within the framework then the framework needs modification.

2.9 Tentative Conclusions

To accommodate the interpreted formalism of quantum field theory, or indeed of any physical theory, any general ontology must be able to provide a home for the concepts of spacetime and its inherent geometrical characteristics, field quantities and other magnitudes, how these are articulated into dimensions and related by invariant equations concerning these magnitudes, their derivatives and integrals. There are deep and difficult philosophical mysteries in all of these matters, and no matter what ontology we use to embed them, these difficulties must be confronted and not shrugged off or ignored. Another issue requiring deep thought is the relationship between these quantities and the mathematical structures we use to represent them; it thus strays into the philosophy of the application of mathematics. Once again this is everyone's problem, and I hasten to stress that in my opinion no one, no philosopher, mathematician, physicist, or hybrid specialist, has come close to a complete solution. Theoretical physicists may become so adept at using mathematical formalism that they may think that the mathematics they employ is a transparent representation. It is not: no representation is wholly transparent, and the facility by which physicists can wield mathematical representations to represent physical situations is in need generally of explication. Because of these difficulties and because of limitations of spacetime I have not ventured into physics or even the philosophy of physics in considering my ontological framework. That was not what, as a journeyman general metaphysician, I considered my brief. To that extent, I have done less than half of the work required to see how metaphysics can meet and accommodate physics. However, on the basis of considering ontological hypotheses, arguments and counterexamples in many areas over a period of more than a quarter of a century I am moderately confident that both the general methodological line and the general type of framework advocated are likely to prove fruitful. If QFT or any other current physics shows clear deficiencies in the scheme, they must be taken on board and the scheme modified, and it is in this spirit that we must await more running to and fro.

References

Armstrong, D. M. (1997): *A World of States of Affairs*. Cambridge: Cambridge University Press.

Bacon, J. (1995): *Universals and Property Instances: The Alphabet of Being*. Oxford: Blackwell.

Campbell, K. (1990): *Abstract Particulars*. Oxford: Blackwell.

Cao, T. Y. (1997): *Conceptual Developments of 20th Century Field Theories.* Cambridge: Cambridge University Press.

Chisholm, R. M. (1976): *Person and Object.* London: Allen and Unwin.

Ingarden, R. (1964): *Time and Modes of Being.* Springfield: Thomas.

Planck, M. (1910): Die Stellung der neueren Physik zur mechanischen Naturanschauung. *Physikalische Abhandlungen und Vorträge*, Brunswick: Vieweg, 1958, Vol. III: 30–46.

Quine, W. V. (1976): Whither Physical Objects? In Cohen, R. S., Feyerabend, P. K., and Wartofsky, M. W. (eds.): *Essays in Memory of Imre Lakatos*. Dordrecht: Reidel, 497–504.

Schrödinger, E. (1950): What is an Elementary Particle? *Endeavour*, 9: 109-116. Reprinted in Schrödinger, E. (1957): *Science, Theory and Man*. London: Allen and Unwin, 193–223.

Simons, P. M. (1991): Faces, Boundaries, and Thin Layers. In: Martinich, A. P. and White, M. J. (eds.): *Certainty and Surface in Epistemology and Philosophical Method. Essays in Honor of Avrum Stroll*. Lewiston, NY: Edwin Mellen Press, 87–99.

Simons, P. M. (1994): Particulars in Particular Clothing. Three Trope Theories of Substance. *Philosophy and Phenomenological Research* 54: 553–576. Reprinted in: Laurence, S. and Macdonald, C. (eds.) (1998), *Contemporary Readings in the Foundations of Metaphysics*. Oxford: Blackwell, 364–384.

Simons, P. M.(1997): Un mondo senza stati di cose. *Discipline filosofiche*, 7: 29-48. Translated as: Eine Welt ohne Sachverhalte in: O. Neumaier (ed.) (2001), *Satz und Sachverhalt*. St. Augustin: Academia, 229–254.

Simons, P. M. (1998a): Farewell to Substance: A Differentiated Leave-taking. *Ratio N.S.* XI: 253–252. Also in: Oderberg, D. S. (ed.). *Form and Matter. Themes in Contemporary Metaphysics*. Oxford: Blackwell, 22–39.

Simons, P. M. (1998b): Metaphysical Systematics: A Lesson from Whitehead. *Erkenntnis*, 48: 377–393.

Simons, P. M. (1999): Die Urfaktoren des Seienden: eine anaximandrische Alternative. In Nida-Rümelin, J. (ed.), *Rationalität, Realismus, Revision*. Munich, Berlin: de Gruyter, 399–406.

Simons, P. M. (2000a): Continuants and Occurrents. *The Aristotelian Society*, Supplementary Volume LXXIV: 78–101.

Simons, P. M. (2000b): Identity through Time and Trope Bundles. *Topoi*, 19: 147–155.

Simons, P. M. (2000c): How to Exist at a Time When You Have no Temporal Parts. *The Monist*, 83: 419–436.

Simons, P. M. (2002): Why Nearly - But Not Quite - Everything is an Entity. In Monnoyer, J. M. (ed.), *Proceedings of the 1999 Grenoble Conference on Australian Metaphysics*. (Forthcoming.)

Simons, P. M. and Ganthaler, H. (1987): Bernard Bolzanos kosmologischer Gottesbeweis. *Philosophia Naturalis*, 24: 469–475.

Simpson, G. G. (1961): *Principles of Animal Taxonomy*. New York: Columbia University Press.

Strawson, P. F. (1959): *Individuals. An Essay in Descriptive Metaphysics*. London: Methuen.

Van Fraassen (1991): *Quantum Mechanics: an Empiricist View*. Oxford: Clarendon Press.

Chapter 3

'Quanta,' Tropes, or Processes: Ontologies for QFT Beyond the Myth of Substance

Johanna Seibt
Aarhus University

During the formational phase of analytical ontology in the first half of the 20th century, ontologists prefaced their investigations with rather extensive methodological considerations. Current research in analytical ontology, on the other hand, pays comparatively little attention to the discipline's methodology. At best there is a quick gloss of ontology as the 'theory of truth-grounds' or of 'truthmakers' without any further elucidation of the data, tools, scope, and purpose of such a theory. But the discipline of ontology hardly can be said to have reached the state of 'maturity' that could justify the neglect of methodological questions. There is no clear shared understanding about what the evaluative criteria for ontologies should be and how considerations from epistemology, metaphysics, and semantics should relate to ontological arguments.[1] More importantly, however, there is no clear shared position concerning legitimate constraints on ontological theory formation, so that necessary conditions of 'entityhood' would be separated from the default assumptions of a particular theoretical tradition. Most contemporary ontologists do not even appear to be aware of the fact that they operate within the confines of a longstanding research paradigm which powerfully restricts the space of solution strategies in ontology.

Without proper attention to current constraints on ontological theory formation, the productivity of the discipline remains significantly limited.

[1] For instance, epistemological concerns about identifiability and reidentifiability have been introduced into the debate about identity and individuation; the metaphysical question about which entities are ultimately real is confused with the question of which entities are ontologically basic; and language-group specific semantic phenomena such as the predominance of sortal nouns in Indoeuropean languages, are ontologically misread as establishing a link between countability and individuality.

Old problems remain unsolved and new domains of application remain foreclosed. Moreover, unreflected theoretical habituations of researchers in ontology are bound to affect philosophers of science who turn to current ontological research to clarify the commitments of a scientific theory. Their interpretive efforts are likely to be fraught with the dispensible presuppositions of a specific research paradigm.

My primary aim in this essay is to assist philosophers of science working on the ontological interpretation of QFT by highlighting some questionable assumptions imported into the debate from the ontological tradition. After some methodological preliminaries, I will state a set of presuppositions, ubiquitous within historical and current ontological research, which are by no means laws of thought but the contingent commitments of an ontological research tradition I call the 'substance paradigm'. Some of these presuppositions, I argue, made their way into the debate about indistinguishability, particlehood, and individuality within the quantum domain. This has the effect, among others, that philosophers of physics misrepresent the case of 'indistinguishable particles' as having decisive bearing on the question of whether 'quantum particles' are individuals or non- individuals. I then consider possible strategies of gaining an ontological interpretation of some core concepts of QFT by modifying the presuppositional depth-structure of the ontological tradition: P. Teller's "quanta" and trope theory, which each operate *with concrete particular entities that are not independent individuals*. In the final section I sketch an ontology based on *concrete individuals that are not particular entities*, the theory of free processes (APT). APT is a genuinely new ontological scheme where a large number of traditional ontological presuppositions are rejected; in consequence it has a variety of features that *prima facie* would seem to allow for a straightforward application in quantum domains. Whether this impression is justified remains to be seen; my primary aim here, to restate, is to argue not for a substantive but rather for a methodological thesis: that the ontological interpretation of QFT must pay critical attention to existing theoretical biases within ontology.

3.1 A Methodological Preface

Methodology always makes for a tedious read. But it is certainly the primary area in which ontologists can contribute to extant research on the ontology of QFT. An ontology of QFT worth the name not only needs to 'get the physics right.' Moreover, a modicum of methodological patience has considerable systematic pay off, as I hope the following sections will show.

(a) *What is ontology?*

Contributions to the ontological interpretation of QFT frequently use "ontology" and "ontological status" in close association with the metaphysical predicate "realist" and "reality." Compare for instance the following quotation:

> [1] The basic ontology of a theory is taken to be the irreducible conceptual element in the logical construction of reality within the theory. In contrast to appearance or epiphenomena, and also opposed to mere heuristic and conventional devices, the basic ontology is concerned with real existence. That is, it is not only objective, but also autonomous in the sense that its existence is not dependent upon anything external, although its existence may be interconnected with other primary entities. As a representation of the deep reality the basic ontology of a theory enjoys great explanatory power: all appearance or phenomena described by the theory can be derived from it as a result of its behavior.[2]

This is in contrast to the rather more reticent and modest attitude of working ontologists who prefer to present an ontology as a theory of truthmakers for a theory T, i.e., as a description of a model structure for T.[3] (T may be a scientific theory or the common sense reasoning embedded in a natural language.) Such descriptions are not, as such, equated with descriptions of reality but serve the purpose to explain or justify the categorial inferences licensed by T. A categorial inference is a subtype of material inferences, deriving from the lexical meanings of the terms involved. More precisely, if G is a 'genus term' of T implied by a number of specific predicates F_i in T, then the categorial inferences licensed by sentences containing

[2]Cao 1999: 4.

[3]This is largely a reconstruction of a methodological stance that is implemented but, as I mentioned, rarely explicitly stated. For an explicit statement compare for instance Fine 1991.

F_i are those that derive from the lexical meaning of G[4]; the patterns of categorial inferences licensed by a genus term G in T are commonly assumed to exhaust the meaning of G.

There are two types of projects involved in developing an ontological description of a model structure for T, which are frequently intertwined. On the one hand, when ontologists announce that they shall 'give an account of' things (or persons, events etc.) they aim to devise a definition for a class of entities (a 'category') in the domain of T whose features ('individuality,' 'concreteness,' 'persistence' etc.) entail all the patterns of categorial inferences licensed by the T-term 'thing' ('person,' 'event' etc.). To establish the relevant entailments, on the other hand, the category features (individuality, concreteness, persistence etc.) themselves need to be defined; so-called 'theories of indviduality' (concreteness, persistence etc.) aim to model *some of* the categorial inferences of *a number* of different G-terms of T (e.g., a theory of concreteness must apply to anything—things, persons, events—to which location expressions apply in T.) In each case the data of an ontological theory are ultimately always the (patterns of) categorial inferences licensed within T. Thus we may say that in its perspicuous methodological form an ontological theory has the form of the quadruple $\langle$ M, O, f_C, S $\rangle$: it specifies an assignment f_C which correlates the elements of a class S of T-sentences with structures of a model M as described by a domain theory O, in such a way that the set of categorial inferences C licensed by S is entailed by the values of f.[5]

(b) *What should be afforded by an ontology of QFT?*

There are four aspects of this model-theoretic reconstruction of ontological research that are particularly noteworthy in the present context. First, it

[4]For a simple illustration, consider the following sentences:

(1) Kim's (only) car is white.
(1) $\Rightarrow$ (2) Whatever is to the left of Kim's car is not Kim's car.
(1) $\not\Rightarrow$(3) Whatever is to the left of Kim's car is not white.
(4) Kim saw the explosion that destroyed her car today at 7am.
(4) $\Rightarrow$ (5) Kim saw something today at 7am.
(6) Kim saw the man that destroyed her car today at 7am.
(6) $\not\Rightarrow$(7) Kim saw something today at 7am.

That (1) implies (2) but not (3), that (4) implies (5) but (6) does not imply (7) derives from the meaning of the genus terms 'thing,' 'property,' 'event,' and 'person' that are embedded in the lexical meaning of 'car,' 'white,' 'explosion' and 'man,' respectively.

[5]I am omitting here a variety of finer points, cf. Seibt 2000b.

highlights the fact that ontological research is *per se* metaphysically neutral. Ontological theories specify what makes T-sentences true without being committed to any particular theory of truth. That ontological research can, and mostly does, remain metaphysically neutral means in particular that the *ontological evaluation* of ontological theories can be *separated from their metaphysical evaluation*. Ontologists, like scientists, evaluate a theory with respect to its *data adequacy*, i.e., with respect to how well it models the relevant set of categorial inferences. The metaphysical status of T itself is not subject to ontological debate. In application to our present context, the question of whether QFT is perhaps not a likely candidate for an 'ultimate theory'–in view of its possible incompatibility with the general theory of relativity—is as such of no concern for its ontological interpretation.[6]

Second, the model-theoretic characterization of ontology as a theory of categorial inferences in T highlights the fact that ontologies are relative to a theory in the sense that they are developed for an *explicitly stated data set of categorial inferences in T*. The ontological interpretation of QFT, however, appears to be guided by commitments that are not part of the explicit inferential space of T. Compare the following passages:

> [2] In a sense, one could say that, at least in QM, the observer sees only massive particles, and only the theorist deals primarily with fields some of which are unobservable. But this does not mean that photons are not particles; few physicists would deny the photon ontological status.[7]

> [3] A strong case can be made that empirically only particles are observed, and fields, except for the classical fields, are not observable. This suggests relegating the concept of the field to the status of a convention, a device for generating particles and mediating their interactions.[8]

Unless operational or positivist intuitions of the practitioners of T are explicitly reflected in the inferential space of core concepts of T—e.g., in terms that contain the categorial inference pattern 'x is unobservable → x

[6]Of course, questions that pertain to the consistency of T are important to the ontological interpretation of T.

[7]Rohrlich 1999: 360.

[8]Cao 1999: 8.

is conventional'—they cannot enter ontological considerations. One might choose to put an additional, *metaphysical* constraint on ontological theory construction by adopting the *metaphysical* tenet that unobservables are conventional. But this is not a common presupposition of ontological theory construction *per se.* Ontological theories—compare ontologies of mathematical entities or mental entities—deal frequently with unobservables.[9]

Third, unless the relevant data set of categorical inferences licensed by T has been fixed or at least roughly delineated, the ontological interpretation of T cannot begin. This explains why the obstacles for an ontological interpretation of QFT are perhaps greater than anywhere else. In order to determine the inferential space of core concepts like 'field' or 'quanta,' the following questions would need to be answered:

(1) Which of the various mathematical representations of QFT—Fock space representation, path integral representation, or algebraic approaches—describe best the physical content of core concepts of QFT and thus should be the primary source in determining the inferential space of these concepts?

(2) How much of a certain mathematical formalism can count as "surplus formal structure" [10] introducing entities that have no causal role?

(3) How much of the formalism can count as 'surplus inferential structure,' i.e., how much of the structure of the representation is relevant for drawing the right inferences about what is represented? E.g., should ontologists pay attention to the operator-eigenstate formalism or should they concentrate on the interpretation of probability amplitudes and superpositions? Does the difference between 'field quantization' and 'second quantization' matter for an interpretation of the concept of quantum field? [11]

[9]In order to capture the spirit of quotations [2] and [3] one would need to extend QFT to contain a sentence class R of experimental results. Then one might argue in the process of data determination that only those concepts should be modeled ontologically which also occur in R (i.e., which are observable). Note, however, that such arguments would nevertheless be based on metaphysical, not ontological grounds.

[10]Redhead 1980.

[11]Cf. Redhead 1988: 15; or perhaps neither is relevant: "Although this is often talked

As far as I can see, there are no uncontroversial answers to any of these questions. The particular weakness of the discussion in section 3.3 through 3.5 of this essay—this must be admitted right from the outset—consists in the fact that I am not capable of offering an explicitly developed data determination for an ontology of QFT but take my bearings from the extant discussion.[12]

Fourth, there is one important characteristic on ontological theories that is not included in the model-theoretic definition as stated so far. An ontology is an *explanatory* theory of categorial inferences in T. This imposes two important requirements on the ontological descriptions of model domains. On the one hand, these descriptions should involve a minimal number of primitive concepts. This *minimality requirement*—well known as 'Occam's razor'—is a common desideratum in explanatory contexts. On the other hand, there is the requirement I call, following Carnap, the *requirement of foundedness* which demands that the basic concepts of an ontology, even though they may be axiomatically defined, are founded in experience. An ontological description is something that we need to 'understand' in the pragmatic, Heideggerian sense of the term in which we do not 'understand' infinite vector spaces and the imaginary parts of complex numbers. For this reason ontologists preferably resort to formal theories with basic generative relations that are well-entrenched in agentive experience, such as mereology or mereo-topology. More involved mathematical formalisms are ontologically suspect in so far as they encode inferential structure and operate with generative relations that are intuitively less accessible. It may be more elegant to resort to the rich inferential texture of, say, sheaves theory, but it is certainly against the primary goal of ontology, namely, to explain or justify the inferences licensed by a theory T by making them structurally trans-

about as second quantization, I would like to urge that this descriptions should be banned from physics, because a quantum field is not a quantized wave function. Certainly the Maxwell field is not the function of the photon, and for reasons that Dirac himself pointed out, the Klein-Gordon fields that we use for pions and Higgs bosons could not be the wave functions of the bosons. In its mature form, the idea of quantum field theory is that quantum fields are the basic ingredients of the universe, and particles are just bundles of energy and momentum of the field. In a relativistic theory the wave function is a functional of these fields, not a function of particle coordinates" (Weinberg 1999: 242).

[12]In particular, Auyang 1995, Bartels 2000, Cao 1999, French and Krause 1999, Kaiser 1999, Kuhlmann 2000, Redhead 1980, 1999, Rohrlich 1999, Teller 1995, 1999, Weinberg 1999.

parent, i.e., by offering structural descriptions of the domain of T which entail the relevant inferences in T in rather obvious ways.

The explanatory requirement of foundedness also accounts for the fact that some ontologies, despite obvious deficiencies, are surprisingly widely accepted and long-lived. The popularity of substance ontology, for example, is not due to its (rather poor) explanatory achievements,[13] but mainly due to the fact that the technical term substance is the categorization of a genus-term of common sense reasoning that we agentively understand particularly well: things. It is the practical 'disclosedness' of things that primarily accounts for the appeal and alleged naturalness of ontological descriptions based on the notion 'substance' or 'object.' Since things—the building blocks and billiard balls of our childhood—are dominant functional elements in our agentive understanding of 'world,' we are strongly disposed to prefer ontological descriptions of theoretical domains that characterize the latter as assemblies of thing-like entities. Competitors of substance ontology thus should better operate with a category that is as well founded as the notion of 'substance' or 'object.'

(c) *What is an ontological explanation?*

Based on the model-theoretic characterization of ontological theories, the explanatory goals of ontology can be clearly determined. The *explananda* of an ontological interpretation of T are its data, the patterns of categorial inferences licensed in T by the core concepts of T (e.g., 'field,' 'particle,' 'quanta,' 'excitation' etc.). The ontological domain theory O provides a structural description of a model M for T by defining basic entities and derivation relations for the formation of complex entities.[14] The *primary explanantia* of an ontological interpretation of T are the 'categories' or basic entity types introduced by O. Categories may be implicity defined by the axioms of O or explicitly defined in terms of category features such as individuality, concreteness, persistence, dynamicity, unification, countability,

[13]Cf. Seibt 1990b.

[14]Cao equates the ontology of T with its "basic ontology" T, a description of "primary entities" from which other entities ("appearances and phenomena") can be derived which do not belong to the ontology of T, cf. Cao 1999:4, quotation [1] above. This deviates strongly from the common understanding of ontological theories whose very purpose it is to give a structural description of the domain which contains definitions for basic entities and derived entities as well as for the derivation relations involved. Cf. for instance Bergmann 1960 or Fine 1991.

determinateness, discreteness etc.[15] The value of an ontological explanation depends on whether category features are well-defined, again either explicitly or, preferably, implicitly. Since in the end it is not categories but category features that carry the data derivation, category features are the *ultimate explanantia* in an ontological explanation.[16] *It is important to realize, however, that ontological explanations cannot leave out the step of specifying categories, relying on single category features only (cf. section 3.3 below).* For category features, which apply to more than one category, are never founded in the sense just described. In our agentive experience we deal with things, processes, states, opportunities etc., i.e., with the denotations of the genus-terms of common sense reasoning, but never with 'genus-transcendent' entities that would correspond to category-transcendent features like individuality, concreteness etc.[17] Of course all sorts of descriptions of the model-structure of T are possible. But, in consequence of the explanatory goals of ontology, only those descriptions qualify as ontological descriptions that involve categories defined in terms of two or more category features, with an additional specification of how entities of this type relate to other entities of basic or complex types.

(d) *Why are there so few ontologies?*

Ontological theories, like any others, are developed within a research paradigm in the Kuhnian sense of a certain set of sociologically entrenched preferences concerning the discipline's problem space, solution space, and methodological standards. Remarkably, however, from Aristotle onwards to the present day, ontological theory formation has been largely governed

[15] To illustrate, x is substance := x is concrete, particular, persistent, independent; x is an attribute := x is abstract, universal, non-dynamic; x is a trope := x is abstract, particular, non-persistent, dependent; x is an occasion := x is concrete, particular, non-persistent, dependent; etc.

[16] To illustrate, relating to footnotes 4 and 14 above: Substance ontologists, for instance, would explain the inference from sentences (1) to (2) by the fact that substances are simply defined to have a unique spatial location at any particular time of their existence. For the substance-ontologist the truth-maker of (1) involves also an attribute which by definition does not have the category feature of unique spatial location; thus the inference to (3) is undercut. In contrast, there are ontologists who claim that the truth-makers of (1) consists of attributes only, some of which are bundled together by a special domain relation 'compresence' which ensures the unique spatial location of whatever is bundled with it.

[17] The fact that common sense reasoning does not have terms for such genus- transcendent features should provide sufficient evidence in support of this claim.

by one and the same research paradigm. Aristotle famously experimented with a category—later translated as 'substance'—to which he assigned a large number of category features. A substance was said to be persistent, the locus of change, countable or one of its kind, concrete, particular, non-instantiable or a logical subject, independent, discrete, simple, unified.[18] This amounts to a functional overdetermination of the category, creating a variety of inconsistencies. The history of substance ontology can be reconstructed as the continued effort of eliminating these inconsistencies by offering functional diversifications of the notion of 'substance,' i.e., by defining 'substance' in terms of various maximal consistent subsets of the Aristotelian feature list.[19] However, these reformative programs in most instances did not abandon the 'theoretical depth-structure' of the Aristotelian approach, a set of presuppositions underlying the Aristotelian assignment of category features. In fact, fairly independently of how the category of 'substance' is defined and even of whether it is retained at all, these presuppositions—here called 'Characteristic Aristotelian Presuppositions'[20] (CAPs)—have acted as constraints on ontological theory formation throughout the history of the discipline. For later reference I list here ten of the twenty or so CAPs (note that the labels of category features–'concrete,' 'individual,' 'particular' etc.–are used here as variables, ranging over the various extant definitions of these features):

Ontological closure clause:

CAP-0: There are at most two types of entities: concrete, individual, particular entities ('substances,' 'objects,' 'particles' etc.) and abstract, universal entities ('attributes,' 'properties,' 'relations' etc.).

Linkage of individuality and countable oneness (numerical identity):

CAP-1: The explanatory factor that accounts for the individuality

[18]Cf. Metaphysics 1042a34, Physics 200b33, Metaphsyics 1038b35f, 1017b16ff, Categories 2a13ff, Metaphysics 1037b1ff, Categories 3b33, Metaphysics 1041a4f, 1041b11ff, respectively.

[19]Cf. Seibt 1990b.

[20]The adjective 'Aristotelian' would doubtless require more detailed historical commentary; suffice it to say that I do not mean to suggest that Aristotle's rich ontological thought, which is as much 'process-ontological' as 'substance-ontological,' is limited to these principles.

of an entity also accounts for X's countable oneness. (Conversely, in so far as 'a' and 'b' can be said to refer to two entities, the denotations of 'a' and 'b' are different individuals.)

Principle of analysis as structural description:

CAP-2: The individuality (thisness) of an entity α is to be explained in terms of a component of the structural description of α (hereafter: the 'individuator of α'). Any qualitative determination (suchness) of an entity α is to be explained in terms of a component of α (hereafter: the 'qualificator of α'; commonly called 'property,' 'attribute' etc.)

Principle of abstract generality:

CAP-3: Qualificators are abstract. Qualificators are not individuated by spacetime location.

Principle of categorial dualism:

CAP-4: Qualificators and the entities they are components of (e.g., 'objects') are to be characterized in terms of mutually exclusive category features.

Linkage between individuality and subjecthood:

CAP-5: The individuator of an entity α is the logical subject of any qualificator of α.

Principle of the ontological significance of linguistic surfaces:

CAP-6: Linguistic structure is a reflection of ontological structure.

The *replica* theory of predication:

CAP-7: An entity α is the logical subject of the (predicate denoting a) qualificator ϕ of α iff α is an example of ϕ (the whole or part of α is like ϕ).

Linkage of subjecthood and independence:

CAP-8: All and only entities which are logical subjects are independent.

Linkage of particularity and individuality:

CAP-9: All and only individuals are particulars.

Principle of determinateness:

CAP-10: All and only entities which are concrete individuals are fully determinate. (Conversely, no concrete individuals are determinable).

The full list of CAPs suffices to define the dominant paradigm in ontology which I call the 'substance paradigm' or—in analogy to presuppositional prefigurations uncovered in epistemology, semantics, or philosophy of mind—the *'myth of substance'*.[21] In the course of the historical hegemony of the substance-ontological tradition, these principles have been so deeply entrenched in ontological research that they are frequently considered as laws of thought. It should be fairly obvious that some of these presuppositions also are involved in the notion of a 'classical particle' in physics and, accordingly, that any attempt to devise interpretations of non-classical or quantum domains would do well to begin with a critical look at these principles.

The quasi-axiomatic characterization of the substance paradigm or the myth of substance I have sketched here immediately suggests a heuristic for the development of ontologies by means of axiom variation or, equivalently, the development of new categories by means of feature combinations that break the established linkages between category features. In fact, this is the heuristic that theory revision within ontology can be shown to have followed implicitly all along. New ontologies—some of which we will consider below—undercut some of the mentioned 'CAPs' and whatever explanatory potential they gain derives directly from the rejection of substance-ontological principles. Since there are a number of alternative definitions for each category feature, there is a rich combinatorial space in which new categories can be placed.

If we were to represent categories as vectors in a discrete vector space or as positions in a matrix, such a space would have about 10 dimensions, corresponding to the number of category features, each with half a dozen values corresponding to the number of alternative definitions for category features. The best approach, both within ontology in general and in particular for the ontological interpretation of QFT, is to think of all positions in this matrix as initially legitimate and to let ontological research decide on—rather than habituation or presumptions about an alleged 'synthetic apriori'—which categories are viable and useful.

3.2 'Identical Particles' and the Myth of Substance

Following this general strategy of category construction by axiom variation, let us investigate then which of the traditional constraints of the substance

[21] For the full list of CAPs cf. Seibt 1990b, 1996a, 1996b, 1996c, 1996d, and 2000c.

paradigm are most likely to be rejected in pursuing the project of devising an ontological interpretation of QFT. The natural starting point for such an inductive approach is what is commonly called 'the traditional notion of substance.' Upon a closer look, this is an empty description since, as I mentioned, diversifying Aristotle's functionally overdetermined notion, the substance-ontological tradition developed at least a dozen different accounts of substance that overlap (in extension and intension) only marginally—there is no substance to 'substance.'[22] But we can discern trends and clusters. The four primary functional roles or category features of 'substance' are independence, subjecthood, persistence, and ultimate determinateness. So we might take each one of these four features and examine how they should be combined with other elements of the list of category features (individuality, concreteness, particularity etc.), to cover the relevant data, i.e., the inference patterns valid in T.

For the benefit of those who hold the notion of a 'classical particle' for ontologically straightforward, it might be useful to begin with pointing out that whatever the correct ontological category of classical particles may be, it cannot be a category that has as category features all three: persistence, subjecthood, and determinateness. This may be surprising since persistence, determinateness, subjecthood (and individuality) are most frequently used to contrast classical particles and 'quantum entities,' and we frequently find ontological classifications like the following:

> [C] The T-sentence 'P is a classical particle' is made true by a concrete, countable, particular, individual, persistent, determinate entity α that has a continuous spacetime trajectory and is the logical subject of T-predicates truly predicated of P.

Upon a closer look, however, it turns out that such an interpretation is hardly recommendable. If we take classical particles to change in the sense of having different properties (e.g., spatial locations) at different times, they cannot at once be said to be (a) the subject of their properties, (b) determinate and (c) persistent in the sense of being strictly identical at different times ("endurance account" of persistence). For assuming [C] the following *aporia* arises.

[22]Cf. Seibt 1990b, concurring with the verdict of Stegmair's (1977) careful study.

(1): Identity, whether applied to entities existing at the same time or at different times, is a relation governed by the Leibniz Law (i.e., by the principle that if $x = y$ then for all f, $f(x) \leftrightarrow f(y)$).
(2): α is the logical subject of predicates that can be truly predicated of P (at some time t).
(3): α is determinate, i.e., at any time t α is the logical subject of any predicate that can be truly predicated of P at t.
(4): P changes during the time period $[t, t']$.
(5): P changes during $[t, t']$ iff F is truly predicated of α at t and not-F is truly predicated of α at t'.
(6): Inference from 1, 4, 5: the α of which F is truly predicated at t is identical with the α of which not-F is truly predicated at t'.
(7): Inference from 3, 6: all predicates truly predicated of the α of which F is truly predicated at t are also truly predicated of the α of which not-F is predicated at t'.
(8): Inference from 5, 7: P does not change.[23]

Even though familiar ordinary things and classical particles may be the entities we are intuitively most familiar with, they are far from being ontologically unproblematic. Given the difficulties ontologists encounter in accounting for the transtemporal identity of classical particles, it is rather ironic that many seem to hanker after ontological correlates of 'particles' also within the quantum domain—as though classical particles were ontologically well-understood and one could expect similar success from a theory of non-classical particles.

Let us turn now to the interpretation of " 'quantum particles,' for want of a neutral term"[24] or, as I shall rather say, 'quantum entities.' Quantum entities have been said to differ from classical particles in that (i) they are not determinate in all their evaluative respects but are essentially de-

[23] There are various versions of this aporia to be found in the literature, e.g. in Haslanger 1989, Seibt 1990b, 1996. This one has the advantage to undercut various apparent rescue strategies that operate with time-indexed properties or time-indexed predications. Note that here no assumptions whatsoever are made about the ontological interpretation of predication, nor whether it is legitimate to time-index the reference to the 'subject' (i.e., use descriptions like α-at-t).

[24] French and Krause 1999: 325.

terminable, (ii) they do not have a continuous trajectory in spacetime or phase space and thus cannot be considered to be persistent entities, (iii) they cannot be individuated in terms of location and momentum. Interestingly, quantum entities share these characteristics with common sense properties like *red, sweet,* or *heavy*, which, in combination with the fact that properties can be multiply occurrent, would naturally suggest an interpretive strategy that associates quantum entities with ontological counterparts of properties.[25] The main reason for why this strategy is, however, rarely pursued is, I believe, due to the substance-ontological prejudice that the ontological counterparts of property talk must be abstract (cf. 'CAP-3' and 'CAP-4' above). Rather, the preferred option is to search for the most conservative modification of the ontological correlate of classical particles, something that is not determinate, persistent, and traceable but otherwise retains the category features of [C] above, such as in [Q1]:

> [Q1]: The QM-sentence 'P is a particle' is made true by a concrete, countable, particular, and individual entity α that is the logical subject of T-predicates truly predicated of P.

But [Q1] also confronts severe difficulties, this time due to a conflict between the features of individuality, particularity and subjecthood. Let us begin with a look at various interpretations of individuality, notably the difference between 'descriptive thisness' (IND-1a, b, c) and 'primitive thisness' (IND-2):

> (IND-1a) Accept the Leibniz Principle of the Identity of Indiscernibles (hereafter: PII) and ground the 'thisness' of an entity in its 'suchness,' where this is taken to be a distinct set of qualificators (e.g., attributes).[26]
> (IND-1b) Accept the PII and ground the thisness of an entity in its suchness, where the latter is taken to be a special individuating qualificator which by definition is had by exactly one entity only (haecceitas).

[25] This idea is elaborated in Schurz 1995.

[26] The questions of what counts as a property in this context and how we are to determine the modality of this statement leads to further alternatives within option (IND-1a). Altogether there are 14 different readings of the PII.

> (IND-1c) Accept the PII and ground the thisness of an entity in its spatio-temporal location or spatio-temporal trajectory.
> (IND-2) Reject the PII and claim that the thisness of an entity cannot be grounded in its suchness but must be assumed as "primitive" or "transcendental".

(IND-2) has received particular attention among philosophers of physics since the common labeling techniques in tensor product Hilbert space seem to require that we attribute primitive thisness to particles. For systems with more than one particle having the same fixed properties and thus being in this sense 'indistinguishable,' the formalism seems to suggest that vectors representing a permutation of the logical subjects of the same properties, e.g., $|a(1)\rangle\,|b(2)\rangle$ and $|b(1)\rangle\,|a(2)\rangle$, represent different physical situations. The dialectics engendered by this observation is well-known. Since Bose or Fermi statistics necessitate that we treat such permutation states as one state, it has been said that quantum physics provides not a strong argument *for* but *against* the primitive thisness interpretation of individuality. On the other hand, proponents of primitive thisness reject this conclusion as premature, claiming that proper attention needs to be payed to the fact that the quantum statistics (either Bose or Fermi) mirrors the restriction to symmetric and anti-symmetric wave functions and thus "can be regarded as arising from dynamical restrictions on the *accessibility* of certain states rather than on their ontological coalescence."[27]

The debate about the primitive thisness of quantum entities is frequently represented as a debate about their individuality:

> [4] What is an individual? For a particle individuation may be provided by some essential 'thisness' that transcends its properties. I will call this 'transcendental individuality' or TI for short. Or we may appeal to spatio-temporal (S-T) continuity of its trajectory. But this means we must be able to individuate spacetime points. ...In QM S-T individuation is not available. So if QM particles are to be treated as individuals then TI must be presumed. Any

[27] Cf. Redhead 1988: 12. For formulations of the various stages in the dialectics cf. e.g., Barnette 1978, French and Redhead 1988, Stöckler 1988, Teller 1995: 20ff.

philosophical arguments against the admissibility of TI will then tell against a particle interpretation of QM.[28]

> [5] As is by now well-known, consideration of the Indistinguishability Postulate which lies at the heart of quantum statistics, suggests two metaphysical possibilities: either 'quantum particles,' for want of a more neutral term, can be regarded as individuals subject to accessibility constraints on the sets of states they can occupy, or, alternatively, they can be regarded as non-individuals in some sense...[29]

Such formulations I find unfortunate in two regards. First, as obvious from our list above, the debate about primitive thisness is not concerned with the issue of individuality *per se*, but merely with *one* of various options for a definition of individuality (not exhausted by 'S-T individuation'). *The belief that arguments against primitive thisness present arguments against individuality can only be motivated on the basis of presupposition CAP-9 of the substance paradigm.* If all and only individuals are particulars, as postulated by CAP-9, then quantum statistics indeed spells trouble for the idea of grounding the individuality of 'quantum particles' in descriptive thisness. An entity is a particular if and only if it has at any particular time t (at which it exists) exactly one (possibly not sharply bounded) spatial location. So-called 'indistinguishable particles' are two entities with possibly different spatial locations that have the same descriptive thisness. That is, quantum statistics implies that descriptive thisness is not sufficiently strong to ensure particularity. In other words, if we define individuality in terms of descriptive thisness, e.g., by (IND-1a), quantum statistics implies that there are individuals that occur multiply, which is excluded by CAP-9. However, as I shall argue below, CAP-9 is a perfectly dispensible principle; already in the macro-physical domain there are multiply occurrent individuals, namely, stuffs and activities.

Second, if the debate about primitive thisness is presented as having some bearing on the "particle interpretation of QM," this suggests that one *could in principle* ascribe primitive thisness to the ontological correlates of 'quantum particles.' But this is not consistently possible, for the following reason. Primitive thisness by definition cannot be explicitly defined. So to

[28]Redhead 1988:10.

[29]French and Krause 1999: 324f.

claim that the ontological correlates of 'quantum particles' have primitive thisness is to claim that they consist of two types of ontological constituents: constituents that ground the truth of true predications about them, and a constituent that grounds their individuality—as the literature has it, 'such-factors' and a 'this-factor.' The postulate of an (undefinable!) this-factor remains entirely empty unless one stipulates in addition that the 'this-factor' of an entity α fulfills the function of being the logical subject of (qualificators denoted by) whatever is truly predicated of α.[30] But now a problem arises. A primitive 'this-factor' must be simple; that is, it neither *is* a such-factor nor *has* such-factors as components. In ontology such simple this-factors are graphically called "bare particulars."[31] But how can anything be bare or simple *and* function as a logical subject of predication, i.e., 'exemplify' or 'satisfy' or 'exhibit' all the features we ascribe to the entity it individuates? Consider the following inconsistent set of claims.

> (1: IND-2) The individuality of an entity α is grounded in the fact that α has a primitive 'this-factor' among its ontological constituents.
> (2: Consequence of IND-2) The 'this-factor' of α is a 'bare particular,' i.e., it neither is a 'such-factor' (qualificator) of α nor has it those as constituents.
> (3: Assumption) The individuator of α is the logical subject of all 'such-factors' (qualificators) ϕ denoted by predicates truly ascribed to α.
> (4: CAP-7) Part or whole of α is an example of ϕ, i.e., either is ϕ or α has a constituent which is ϕ.[32]

[30] This double function of 'this-factors' is frequently supported by a more or less conscious exploitation of the fact the indexical 'this' can be used not only as a modifier—as in: 'this chair'—but also independently—as in: 'this is green'.

[31] Some philosophers of physics muddle the waters by lumping 'haecceitas'—an individualizing such-factor—together with 'bare particulars,' cf. e.g. French/Krause 1999: 325f. Whatever one chooses to term 'primitive thisness,' the distinction between individualizing such-factors (qualificators) and individualizing this-factors should be preserved.

[32] This argument is well-known in the literature, cf. e.g. Sellars' pithy formulation: "Perhaps the neatest way in which to expose the absurdity of the notion of bare particulars is to show that the sentence, 'Universals are exemplified by bare particulars,' is a self-contradiction. As a matter of fact, the self-contradictory character of this sentence becomes evident the moment we translate it into the symbolism of *Principia Mathematica*. It becomes, '(x).(Ef) fx.→¬(Eg)gx' ..." (Sellars 1965: 282).

There are only two serious rescue strategies to shelter the primitive thisness account against this inconsistency. The first introduces the—certainly inspiring—distinction between the "naked" and the "nude," which compromises the bareness or simplicity of this-factors. The second and more respectable path is to modify CAP-7 and to offer a non-pictorial account of predication which allows for simples to be logical subjects. Upon a closer look, however, either path produces either a regress or a contradiction.[33]

In ontology it is difficult to develop 'knock-down-drag-out' arguments, but the case against primitive thisness comes very close indeed. To repeat, if 'primitive thisness' is to denote an ontological solution to the problem of individuation rather than merely to label the problem ('particles are individuated by individuators'), then the stipulated bare individuators must fulfill at least one other explanatory function; but, as here instantiated with the function of logical subjecthood that is traditionally assigned to bare individuators, due to their bareness they cannot fulfill any other explanatory function. Thus, it appears that those who have argued for primitive thisness based on 'indistinguishable particles' have been barking up the wrong ontological tree.

3.3 Non-Individual Particulars or Quanta

As I argued in the previous section, it is only on the basis of presupposition CAP-9, linking individuality and particularity, that the occurrence of quantum states with 'indistinguishable particles' can be taken to provide a case against individuality. Only if one is prepared to exclude with CAP-9

[33]In Seibt 1990b I offer a detailed discussion of these two strategies, as proposed by Baker 1967 and Bergmann 1967, respectively. Here are two quick pointers. Nudity first. Nude particulars have properties but no essential properties. If everything which has a property, has a particular among its ontological constituents that is the logical subject of these properties, then nude particulars themselves would seem to have particulars as constituents and so forth. Second, if we change CAP-7 into a non-pictorial or pure constituent account and postulate that to be the logical subject of a property F is to be a constituent of an entity of which F is a constituent, then we either find ourselves again stipulating bare particulars that are themselves complex (regress) or stipulating bare particulars that do not even have the property of being concrete, or simple, or being a bare particular (contradiction). As far as I cansee, the pictorial and the pure constituent account of predication are the only available strategies for an ontological account of predication—the common model-theoretic treatment of predication in terms of the 'satisfaction' relation is ontologically empty.

that individuals may be multiply occurrent or non-particular, one could argue, pointing to the occurrence of 'indistinguishable particles,' that descriptive thisness is too weak to ensure particularity. Even though below I will argue for the dispensibility of CAP-9, let us for a while go along with it and consider various proposals which (based on CAP-9 and the occurrence of 'indistinguishable particles') categorize 'quantum particles' as non-individual particulars.

There are a number of physicists and philosophers who have entertained the idea that 'quantum particles' are non-individuals, approaching this task along two different routes.[34] The first route consists in devising mathematical frameworks as a semantics for Schrödinger logics. For instance, supplementing the axioms of classical set theory (ZF) to allow for non-individual atoms in set-theoretical constructions, French and Krause have presented a "logic of quanta" in terms of a theory of quasi-sets:

> [6] In quasi-set theory, the presence of two sorts of atoms (*Ur elemente*) termed m-atoms and M-atoms is allowed, but the concept of identity (on the standard grounds) is restricted to the M-atoms only. Concerning the m-atoms, a weaker 'relation of indistinguishability' is used instead of identity. Since the latter ... cannot be applied to the m-atoms, there is a precise sense in saying that they can be indistinguishable without being identical.[35]

In other words, 'quantum particles' or 'quanta' are simply characterized as entities to which "classical identity" does not apply.[36] While the

[34]Cf. Teller 1995 and French and Krause 1999, referring to Born (1943), Weyl (1949), Schrödinger (1957), Post (1963), Dalla Chiara and Toraldo di Francia (1993, 1995).

[35]French and Krause 1999: 327f.

[36]French and Krause op.cit. do not explicitly state definitions of the most relevant concepts, namely, of the equivalence relation of "indistinguishability" and of identity for M-atoms. There is some indication that by "classical identity" the authors refer to the "standard" definition of identity in "first order predicate logic with identity" (327). This is puzzling, however, According to the standard definition, identity is determined by the Leibniz Law (indistinguishability of identicals), the principle of identity (reflexivity) and symmetry. Thus, the standard definition of identity in predicate logic is *neutral* vis-à-vis the ontological alternative of primitive or descriptive thisness (cf. IND-2 and IND-1a above). As long as it remains unclear in which way the equivalence relation of 'indistinguishability' differs from descriptive thisness (IND-1a), it is difficult to see how m-atoms could be said to be indistinguishable yet not identical.

constructional principles of such formal semantic frameworks are most significant for an ontological description of 'quanta,' it important to note, however, that offering a characterization of 'quanta' as non-individuals is only a first step in this direction. As I pointed out at the end of section 3.1 above, in order for 'quanta' to qualify as a category label, entities of this type need to be assigned additional category features (e.g., particularity and concreteness).

The other route to 'quanta' seeks to "arrive at the new conception by starting with our preconceptions about particles and eliminating aspects often thought to be included."[37] Following this approach Paul Teller has recently introduced a contrast between "quanta, understood as entities that can be (merely) aggregated, as opposed to particles, which can be labeled, counted, and thought of as switched."[38] In contrast to the first route, this 'eliminative' procedure of selective removal of some of a multiplicity of category features from traditional categories does provide an ontological interpretation of the QFT notion of a field quantum; Teller's 'quanta' form a category: they are non-individuals, but also particular, concrete, and occur only in discrete units.[39]

As common in ontological research, Teller employs "analogies" and "metaphors to help refine the conceptualization." However, there is no reason why the 'eliminative' procedure should *exclusively* rely on the usage of analogies and metaphors. This harbours certain dangers—without supplementary definitions and clarifications the scope and focus of the intended comparison is difficult to control. The crucial conceptual element in Teller's approach is the contrast between the plural occurrence of quanta and the plural occurrence of particles. This contrast is expressed in terms of the difference between "aggregating" and "counting." However, only for the latter predicate we are given some direct commentary. For Teller all and only "old-fashioned particles" are individuals whose individuality is established by primitive thisness and all and only "things with primitive thisness can be counted," where "[in] counting out a number of particles, there is always a difference in principle in the order in which they are listed."[40] How

[37]Cf. Teller 1995: 29ff.

[38]Teller 1995: 37.

[39]Ibd. 30.

[40]Ibid. 29.

the 'aggregation' of non-individuals differs from such counting Teller illuminates with two analogies. (a) For two pennies locked up in a safe-deposit box, one can "ask intelligibly" which one was deposited first and whether they have switched location; for two dollars on a bank account "similar questions make no sense." (b) If two people hold a rope at opposite ends and shake it to create each a "traveling bump,"

> [6] ... we see two bumps continuing down the rope in opposite directions. But does it make any sense to think of a situation in which the two bumps are switched? Could there be such a situation somehow distinct from the original? Hardly! And although initially one may be inclined to think of the two bumps as passing through each other after merging in the middle, it makes just as good sense to think of them as bouncing off each other.[41]

As evident from Teller's own formulations, it is questionable whether the second analogy is sufficiently analogous to the first (originally formulated by Schrödinger) to count as "another example of these ideas." In the rope analogy we cannot decide whether the bumps run through each other or have bounced off from each other, based on what "we see." Even if we mean by the latter: based on what we could ever empirically determine about the rope, the difficulty is epistemological rather than conceptual. If either interpretation "makes just as good sense" as the other, then, in particular, it does make sense to think of bumps on a rope as switched, as traveling from left to right or *vice versa*, as arriving first or second at the left $\frac{1}{3}$ mark on the rope, and so on. What the rope analogy illustrates is merely indistinguishability in the traditional epistemic sense—the feature of individuality can be meaningfully applied to bumps on a rope, even though we may sometimes not be able to determine which individual bump is which. What the rope analogy lacks is an argument why we cannot, in principle, ascribe primitive thisness to bumps on a rope (assuming, for the sake of the argument, that we can ascribe primitive thisness to pennies). Consider the penny version of the rope analogy. Imagine two pennies being put in a safe-deposit which is so hot insider that the two pennies melt; the melted copper is collected in the middle of the box, then directed into two one-penny molds and cooled down. As in the rope case, before and after

[41] Ibd. 29.

the spatial 'merger' we have two spatially locatable particulars for which it does make sense to ask which one is which (they are distinct already in terms of descriptive thisness, thus–proponents of primitive thisness would say—a fortiori by primitive thisness). Only during the time period when the two bumps or pennies are spatially 'merged' or superposed, it does not make sense to ask 'which one is which' in the same (or sufficiently similar) way in which it does not make sense to ask of two dollars on a bank account 'which one is which.'

The point of the money analogy is to provide an illustration for the claim that one might meaningfully ask 'how many?' for entities for which one cannot meaningfully ask "which is which?" The quantized version of the rope analogy equally fails to communicate this point. Consider the continuation of [6]:

> [7] One moves from the classical to the quantum edition of this idea by imposing the requirement that the bumps can come only in sizes of one, two, three or more equisized units, while being careful to retain the condition that the bumps, restricted to discrete units though they be, still do not bear labels of other symptoms of primitive thisness. It makes no sense to think of switching the 'chunks' around. There is no such thing as switching, even conceptually , which would result in a different situation.[42]

Contrary to [6], here Teller explicitly claims that it "makes no sense to think of switching," but we are not given any further argument in substantiation of this claim. As Teller himself stresses, whether the amplitude of the bumps is 'quantized' or not has no bearing on the issue of whether the bumps can be thought of as individuals. Thus our observations concerning [6] reapply to the quantized case—it does make sense to ask 'which one is which' for each bump, 'quantized' in "size" or not, before and after they are superposed in the middle. In order to adjust the rope analogy, so that it could count as another illustration of the point of the money analogy, one needs to focus on that time period of the 'spatial merger' where one has two units (amounts) of what pennies or bumps are made off (copper and momentum). This shift in focus can be performed for 'quantized' bumps. *Pace* Teller's explicit formulation, a bump travelling from left to right, i.e.

[42] Ibd. 30

with a distinct position in visual or coordinate space, arguably qualifies as an individual (based already on descriptive thisness alone). But the same cannot be said of the discrete units of size that make up the size of the bump. Here indeed the parallel to the money analogy is rather obvious—we cannot imagine any spatial or conceptual distinctness for unit sizes of money in a bank account, nor for unit sizes of amplitudes. For unit sizes of amplitudes it does not make sense to ask which is which or whether they have been switched, not because a switch would not result "in a different situation," but because the n component unit sizes of an amplitude of size n never occur in different spatial locations, which is a precondition for the meaningfulness of these questions.

So I think that Teller's rope analogy must be read against the text, focusing on unit sizes of 'bumps' rather than on 'bumps,' in order to render it sufficiently analogous to the money analogy. It is important to see, however, that the rope analogy is more than a rhetorical repetition. The unit sizes of money in a bank account cannot be ascribed *any* spatio-temporal location. In contrast, the unit sizes of the amplitude of an oscillation 'bump' on a rope can be ascribed spatio-temporal location—they are where the 'bump' is. Thus the rope analogy provides an additional category feature to the new category of quanta: they are not only non-individuals but also concrete (spatio-temporally localizable) and particular (spatio-temporally localizable in one—possible vaguely bounded–region).

Let us turn then to the contrast between 'aggregating' and 'counting.' To requote, for Teller all and only "things with primitive thisness can be counted" and counting always requires "always a difference in principle in the order in which they are listed." This suggests that 'aggregable' is used as a metaphor to denote everything which is not countable in the specified literal sense, i.e., the predicate 'is an aggregable entity' denotes the complement set of the denotation of 'is a countable entity.' But I wonder whether such a wide notion of 'aggregables,' which would include non-discrete or abstract entities, really serve Teller's purposes. There are finer distinctions to be drawn—counting does not always presuppose primitive thisness and non-counting also takes a variety of forms—in terms of which quanta can receive a more precise characterization. Consider the following predicates:

(1) *Continuous amassing:* x can be continuously amassed iff x is of kind K and for any y of kind K, if z is the mereological sum of x

and y, then z is of kind K.

(2) *Discrete amassing or measuring:* x can be discretely amassed or measured iff x is of kind K and x is part of S which is coextensive with a sum of non-overlapping parts p_i of kind K, and for any y, z, if y and z are part of S, then also the sum of y and z is part of S.

(3) *Aggregating:* x can be aggregated iff x can be discretely amassed and for any sum S_i, of entities of kind K there is exactly one S_k such that S_k is coextensive with S_i and S_k is a sum of non-overlapping K-'atoms' (i.e., parts of kind K which do not have parts of kind K).

(4) *Cardinal counting:* x can be 'cardinally counted' iff x is aggregable and there is a function f: $\mathrm{P} \rightarrow \mathrm{N}$ from K-atoms in x into the natural numbers.

(5) *Ordinal counting:* x can be 'ordinally counted' iff x can be cardinally counted and any two K-atoms in x are distinct from each other.

To provide some illustrations, when we count the people entering a building we assign them places in an ordered fashion ("the first, the second...") and this certainly requires that they are individuals, i.e., distinct from each other either by primitive or descriptive thisness (cf. 5.). But when we count the marbles in an unopened box by measuring the weight of the box, all that is required is that the marbles be particulars, i.e, that they not be multiply occurrent at any time t of their existence. Teller's intentions in characterizing quanta as 'aggregable' are, as far as I can see, best captured in terms of the predicate of 'cardinal-countable' as defined here.[43] This allows us to reserve 'aggregable', in accordance with common usage in ontology, to denote unstructured groups of discrete particulars (cf. 3.). A pulk of football fans, for instance, is aggregable—the sum of *any* two pulks of football fans, small or large, is again a pulk of football fans. In contrast, structured units of particulars, such as a football team, or anything that comes in a 'quantized size' are commonly *not* said to be aggregable. Teller speaks of "an 'amount' of quanta,"[44] but strictly speaking 'amounts' come

[43] After the presentation of this talk in October 1999 I discovered that in French and Krause 1999 a similar suggestion is made. French and Krause note that the notion of counting is "unclear" and propose to formally represent Teller's distinction between aggregable and countable by the difference between cardinality and ordinality of quasi-sets (cf. 1999: 330).

[44] Ibid.

into play only if we turn away from groups of particulars to *stuffs*. When we count the liters of wine in a barrel by weighing the barrel, we exploit the fact that wine is that kind of stuff that can be distributed into discrete quantities in the sense of [2]. This is not the case for functionally more complex stuffs like fun, charm, luck, creativity, youth, or thoughtfulness of which we can only say that there is much or little or more here than there, but for which there are no discrete measuring units into which they could be divided up even conceptually (cf. 1.).

To sum up, the 'eliminative' strategy of arriving at a new category of 'quanta' yields a category proper. But in stripping off category features from traditional particulars the primary task is to get clear on what is left over, not only on what is to be stripped off. Predicates [1] through [5] can show, I believe, that Teller's contrast between countable and merely "aggregable" entities is less helpful for this purpose than the contrast between ordinal-counting and cardinal-counting as stated. If quanta are said to be "aggregable" in Teller's sense of being not-countable (that is, as an umbrella-concept covering predicates [1] through [4] above), one has effectively stripped the category for classical particles not only of the feature of countability, but also of particularity and discreteness. But Tellerian quanta are supposed to possess particularity and discreteness. In contrast, the predicate 'cardinal-countable' does entail particularity and discreteness (cf. the notion of a K-atom). There may be other and better ways to calibrate Teller's notion of "aggregable" non-individual particulars; I merely want to urge here some conceptual supplementation of the lead metaphor.

3.4 Fields of Trope Structures?

There are many ways to break the spell of the Myth of Substance. In the previous section we considered the strategy of abandoning the traditional linkage between particularity and individuality (cf. CAP-0, CAP-9). Given the prevailing particularist climate created by the Myth of Substance, this option is certainly attractive. Even if Teller's "quanta," i. e. the ontological counterparts for the QFT notion of 'quanta', are not individuals and not ordinal-countable, they are nevertheless, quite in line with the substance-ontological tradition, concrete, discrete, persistent particulars which are also the logical subjects of predicates truly ascribable to quanta. Teller's "quanta" are traditional particulars 'without the attitude.'

The traditional particularist bias is also retained in a second type of revisionary effort which takes off from a critical reflection of the traditional preconceptions about qualificators (attributes, suchnesses etc.). Proponents of so-called 'trope ontology' abandon the traditional principle that qualificators are general entities (conversely, that all and only logical subjects are particulars, cf. CAP-0, CAP-5, CAP-9) and that all general entities are abstract (CAP-3). Rather, a qualificator, the ontological counterpart of a T-predicate, is taken to be a particular: a "trope." Tropes are individuals–there may be two exactly similar tropes t_1 and t_2 (the *red* in this ball vs. the *red* in that ball) whose distinctness is a matter of primitive or descriptive thisness (spatio-temporal location). But tropes are not independent like the individuals of the substance-ontological tradition: tropes come in similarity classes to ground ontologically our predicative comparisons (traditionally the job of 'universals') and they combine—some say necessarily—to form complex tropes some of which are the ontological counterparts of what common sense or science conceive of as things, particles, organisms, people, events etc. In other words, trope ontologies are typically "reductive" ontologies, aiming to replace traditional categories (substance, attribute, state of affairs etc.) by various types of trope structures. The fact that trope structures are conceived not as additions but as replacements of traditional categories accounts for two important potential assets of a trope approach. First, trope ontology may offer new ways to treat or sidestep some central traditional problems without exacting too much conceptual expenditure—tropes are countable individuals and thus should seem familiar enough. Second, and this is of particular importance for the present context, trope ontology appears to be sufficiently versatile to make a suitable candidate for an *integrated* ontological framework able to accommodate the categorial inferences of common sense reasoning and scientific theories.[45]

One of the best reasons for pursuing a trope-based ontology of QFT may be a practical one: the major building blocks of such an approach are already developed and only need to be brought in contact.[46] Thomas Mormann has recently developed the—to my knowledge—mathematically richest version of a trope theory, using the theory of sheaves or topological

[45]Cf. How one might extend a trope-based identity theory for things to microphysical entities is sketched, for example, in Simons 1994.

[46]This remark here is probably superseded by Kuhlmann's most recent work on a trope-ontological interpretation of QFT, cf. Kuhlmann 2001.

fiber bundles.[47] Mormann's approach is primarily designed to solve the problem of universals by defining "global properties" or "'ersatz' universals" as sections of the trope sheaf. The base space of that sheaf is the set of things, i. e., equivalence classes of compresent tropes. As far as I can see, with proper adjustment to a relativistic setting, a trope-ontological framework of this kind could be combined rather straightforwardly with general constructional ideas for an ontological interpretation of QFT recently developed by Sunny Auyang.[48] Auyang herself champions events as basic ontological entities; an event is what happens in a field at a certain spacetime point—roughly speaking, an event is determined by the state the field is in at a certain spacetime location. More precisely, since a state is always given under a certain representation, an event is defined as that which is invariant under transformations ("the local symmetry group") on (items in the event's "quality space" containing the possible states of the field at that point, i. e., on) the possible representations or "possible characterstics" of the event.[49] To ensure that the transformations on the representations of an event are strictly local to that event, Auyang resorts to a fiber bundle approach, with the base space being the spacetime manifold and the fibres being replicas of the quality space. In associating two spaces, the quality space determining what kind of an event is taking place, and the base-space of the spacetime manifold determining the identity of an event, the fiber bundle approach "highlights the idea that what we take as objective features [i. e, events] are actually a product of a qualitative 'dimension' and an identifying 'dimension.'" But in order to avoid substantival spacetime Auyang also stresses that bundles can be taken as ontologically primary and defines spacetime as the quotient of the bundle by the equivalence relation of 'being in the same fiber,' defined on characteristics of events (hereafter: relation SF).[50] Pictorially speaking, spacetime is a 'cut' through a bundle of 'coincident', i. e, SF-equivalent, possible characteristics of events. However, this move saddles the approach with a dilemma. (a) Either spacetime points have the function to individuate events, as Auyang explicitly an-

[47] Cf. Mormann 1995.

[48] Cf. Auyang 1995.

[49] Cf. ibid. 216.

[50] I am omitting here structural properties of spacetime (defined in terms of coordinate representations invariant under the spatio-temporal group) to simplify the presentation of the argument.

nounces characterizing spacetime as a "divider."[51] But then a circularity threatens. For the definition of SF refers to items occurring in the same fiber, thus to representations or possible characteristics of *some individual* event. (b) Alternatively, if we start with events already introduced as individual particulars, there are two routes to go. (b1) First option: We assign to events primitive thisness. But then we encounter the difficulty set out in section 3.2 above that such talk about primitive thisness cannot consistently be cashed out within a compositional analysis of events but remains a mysterious postulate. (b2) Second option: We assume, in a manner reminiscient of Russell's procedure in *Human Knowledge*, that there is always sufficient qualitative difference between what happens in different spacetime locations so that individual point events or spacetime points can simply be defined in terms of their descriptive thisness. I am not clear on whether (b2) is really an option, however. Can one reasonably assume that for any spacetime locations x, y, and any sets of possible states of the field $S1$ and $S2$, if $S1$ is associated with x and $S2$ is associated with y, then $S1 \neq S2$? If this question cannot be answered 'yes,' the best way to resolve the quandary would seem to be to turn to trope theory. If the replicas of the quality space constituting the fibres in a fibre bundle are taken to represent tropes, then SF can be defined on individual tropes instead on events and the circularity is avoided.

In short, Auyang emphasizes, in the spirit of a traditional substance-ontological dualism (cf. principles CAP-0 and CAP-4 above), that basic ontological entities are definitional constructs of a qualifying and an individuating 'dimension'; the qualifying dimension consists of general entities, the individuating dimension of particulars (spacetime points). She also insists on the reducibility of the individuating dimension which creates a curious tension in the approach; on pain of circularity, the scheme is committed to 'constructing' individual particulars (spacetime points) out of general qualificators (characteristics of events conceived of as universals). Trope ontologists on the other hand begin with individual, particular qualificators which make the derivation of any type of individual particular entity (spacetime point or event or object) comparatively straighforward. Instead of playing the traditional game of deriving individuators from qualificators or *vice versa*, tropists start with entities which establish both the individuating and the qualifying 'dimension' at once, being a 'this-such.' I think it

[51] Ibid. 139.

would be an interesting project to try and rework Auyang's interpretation of QFT within a sophisticated trope-ontological framework along the lines of Mormann's approach; this would afford us not only with Auyang's intended notion of spacetime as an individuating yet derived 'dimension,' but also with ways to define ontological counterparts for quanta, for instance as locally restricted trope sheave sections that fulfill the predicate of being cardinal-countable.

But before engaging in such a project trope ontology needs to be defended against a variety of fundamental objections. In a nutshell, it has been charged that the very notion of a trope is conceptually incoherent—relative to extant expositions, there is no way to make coherent sense of the individuality, dependence, and particularity of tropes.

3.5 Fields as Free Processes?

Somewhat ironically, the only contemporary ontological scheme that has received serious attention from philosophers of physics is one that most ontologists are not familiar with. Alfred N. Whitehead's "philosophy of organism" is explicitly developed as an integrated ontology, aiming to capture entities of all energy scales. It is the only complete metaphysical system developed in 20th century analytical philosophy and its explanatory power—if we take that to sum explanatory compass, systematic depth, and semantic precision—is perhaps unique in the history of metaphysics altogether. For reasons of space I can only recommend the extant research on a Whiteheadian interpretation of QFT to the reader's attention.[52]

The previous two sections dealt with candidate ontologies for QFT which retain the substance-ontological presupposition that the world is an assembly of particulars, whether individual ('substance'), non-individual ('quanta'), or dependent ('trope'). Also a Whiteheadian 'occasions' are particulars. I want to conclude now with a brief look at an ontology that rejects the particularist stance altogether, taking off from the denial of the substance-ontological presupposition (CAP-0, 3, 4, 9) that all concrete individuals are particulars.[53]

[52]Cf. e.g., Shimony 1965, Folse 1982, Stapp 1977, 1979, Malin 1988, and, in particular, Hättich 2000.

[53]To my knowledge, the only other place where this traditional assumption is questioned is Zemach 1970, where a category of concrete, multiply occurent individuals called

"Axiomatic Process Theory" (APT) deviates even more profoundly from the substance paradigm than Whitehead's theory of occasions. However, unlike Whitehead's scheme it is terminologically parsimonious and takes its bearing from common sense reasoning about 'stuffs' and 'activities'—genera terms whose ontological analysis until recently has been all but neglected in the substance paradigm. The notion of a 'free process', the basic entity in APT, is gained along the following line of observations. Some ontologists believe, and that is certainly the crudest manifestation of the myth of substance, that all and only individuals are thing-like entities or substances. But there are many individuals which are countable yet they are not things—a smile, a bunch of grapes, the pace of a city, a vortex, a wedding, your center of gravity, an opportunity, your voice. Moreover, there are many individuals that are not even countable or 'one of a kind.' However individuality may be defined, reidentifiability in the sense of (P1) is certainly an uncontroversial sufficient condition or criterion for individuality:[54]

> (P1) If 'the A which ϕ is the same as the A which χ' can be truly said in T, then the denotation of 'the A' is an individual.

According to (P1) some of the individuals we talk about are stuffs: entities like water or music which do not come 'pre-packaged' in countable units but can be merely continuously amassed or measured in the sense of definitions (1) and (2) in section 3.4 above.[55] We also use nouns for activities ('work,' 'sport,' 'weather') in the sentential context specified in (P1). Activities are

'types' is introduced.

[54] I. e. for a notion of individuality that is not immediately tied, by substance-ontological presuppositions, to concreteness, particularity, ultimate determinateness etc.

[55] Compare: 'The same wine can taste rather different in different glasses'; 'this is the wood we had in our kitchen before the house burned down'; 'the music he was listening to in the car was the same I heard in the radio at home' etc. Most readers will be tempted to interpret such statements as shorthand for talk about relationships between instances of the same kind of wine, wood, or music. But this is not what we say and such a reading only reflects how our intuitive understanding of natural language and common sense reasoning is already infused with the complex traditional prejudice I called the Myth of Substance. Turning the tables, we would certainly, and justifiedly, resist the suggestion that a sentence like 'The chair you are sitting on now and the chair I used a long time ago are the same' is just shorthand for a sentence about the relationships between amounts of dynamic stuffs or activities: 'what you are sitting on now is the same amount of chairing I used a long time ago.'

thus, like stuffs, individuals; they are also, like stuffs, concrete and non-countable. That is, like stuffs activities are homomerous or like-parted: (almost) every spatial or temporal part of a region in which there is wood, music, working, or snowing, is a region in which there is wood, music, working or snowing of the same kind, respectively. In contrast, nouns for things and events (structured developments) are connected with the categorial inferences signaling countability, particularity, and inhomomerity—roughly, no spatial part of a table is a table (of the same kind) and no temporal part of an explosion is an explosion (of the same kind). In fact, the categorial inferences governed by nouns for stuffs and activities on the one hand, and nouns for things and events on the other hand, are sufficiently similar to devise a common ontological interpretation for each pair: (spatio-temporally n-dimensional) non-countable or homomerous entities and (n-dimensional) countable or inhomomerous entities. Further, countability or inhomomerity can be treated as a special case of homomerity:

> (P2) An n-dimensional non-countable entity x is *minimally homomerous* iff for some n-dimensional region R in which there is x, there is no proper part of R which is there is x.

In this way *countables can be treated as a species of non-countables.* Finally, consider assumptions (A1) and (A2):

> (A1) Any $n < 4$-dimensional non-countable entity[56] can be defined as the emergent product of the interaction of 4-dimensional non-countables. (A2) Any $n < 4$-dimensional countable entity[57] can be defined as the emergent product of the interaction of 4-dimensional non-countables.

If (A1) and (A2) can be established by suitable reductive definitions, we have gained a monocategoreal ontological scheme based on 4-dimensional non-countables. Assumptions (A1) and (A2) are the primary research hypotheses of the process-ontological scheme APT based on 4-dimensional non-countables, called 'free processes.'

[56] E. g., the ontological counterparts of talk about stuffs or phenomenal qualities. Note that the dimensionality of an entity is determined by its identity conditions; an entity might exist in time, and have components extended in time, without being itself extended in time.

[57] E. g., the ontological counterparts of talk about things, holes, extensions, shapes, boundaries, or durations.

Like any ontological category, free processes are theoretical entities with axiomatic characterization. But (like substances and unlike tropes and occasions) they are well-founded, i. e., there are inferentially well-embedded common sense genera-terms such as 'activity,' 'happening,' 'goings-on' (or in German: 'Geschehen,' 'Treiben') which indicate that we 'agentively understand' the non-developmental dynamics signified by the technical term 'free process.' Many activities, such as running or reading or rolling, are performed or suffered by animate or inanimate agents (or groups thereof). But not any happening is a 'change in a subject.' Sentences about happenings are frequently not answers to the question what is happening with whom: 'it is raining,' 'it is snowing,' 'it is itching,' 'it is burning.' Even if we express goings-on with sentences with proper noun phrases (compare: 'daylight is coming,' 'the fog is growing thicker,' 'the wind is blowing,' 'the fire is flickering,' 'the vortex is travelling to the right,' 'life its getting easier,' 'the stock market is collapsing' etc.), there is no thing or person (or group thereof) which can be said to do or undergo what the verb expresses.

The notion of a free process is a close cognate of C. D. Broad's and W. Sellars' "subjectless," "absolute," or "pure processes," but there are several decisive differences to note. First, free processes are *not particulars.* They are individuated in terms of their descriptive thisness, not by spacetime location, and may occur in a multiply disconnected spatio-temporal region with fuzzy boundaries. Their spatio-temporal location may even be indeterminate. For, second, free processes are *not necessarily fully determinate*. Just as there are more or less specific stuffs and common activities, so free processes occur in different degrees of determinateness. Where Perrier is, is water and liquid, where sprinting is, is running and exercising—apart from the deep-seated prejudice that the 'building blocks of the world' must be fully determinate in all their respects while generality or indeterminateness can only be a product of abstraction, there is not reason to deny equal concreteness (occurrence in spacetime) to determinable and determinate entities. Third, *simple* free processes are repetitive or dynamically homogenous occurrences *without developmental structure* or culmination point. Fourth, congruent with the syntactic role of nouns for non-countables which are used both in subject and predicate position (cf. 'Perrier is water'—'water is wet'; 'jogging is a kind of running'—'running is relaxing'), free processes undercut the traditional dichotomy of independent logical subjects and dependent qualificators. In sum:

> (P3) Free processes are (i) concrete or spatio-temporally occurrent (ii) individuals that are (iii) 'dynamic stuffs' rather than changes in a subject. (iv) They are non-particulars or (contingently) multiply occurrent. (v) They are not fully determinate, i. e., they have different degrees of specificity or determinateness. (vi) Simple free processes are not directed developments (events) but are dynamically homomerous.

Thus, free processes are *not* particular occurrences with a fixed whence and whither and developmental structure as the term 'process' can suggest in certain contexts. They are not modelled on a single movement of a classical particle with determinate trajectory but on a dynamic conditioning of a spacetime region such as snowing or music. (Free processes may, however, combine to form interaction products that fulfill the logical role of particular developmental occurrences (events), as I shall sketch below.)

APT is based on a non-classical mereology with a non-transitive part relation '$-<$.'[58] Due to the non-transitivity of '$-<$' the so-called proper parts principle can be retained, i. e., without incurring the usual problems with intensionality free processes can be individuated by their 'descriptive thisness,' here total collections of parts. In terms of '$-<$' a large variety of different types of complex processes are defined, some of which are themselves combinations of processes. There are various types of such combiner processes; some of them generate components in the combination product, some suppress components, accommodating phenomena discussed under the headings of 'emergence' or 'complexity.' Due to the non-transitivity of '$-<$' the typology of complex processes is quite rich. The relation '$-<$' is diversified into spatiotemporal, material, and various types of functional part relations, some of which are fully transitive, while others are transitive only to a specified degree. In terms of these relations, and various types of mereological sum and product defined on them, free processes are further

[58]This is to be read as 'is part of' instead of the frequent 'is *a* part of' which introduces surreptitiously a restriction to countable parts. There is an interesting parallel here. French and Krause 1999 point out that set-theory is restricted to individual 'Ur-elements.' Analogously, I have argued (1990b, 1996, 2000c, 2001) that classical mereology contains the rarely reflected presupposition that parts are countable entities; in effect, classical mereology—whose basic relations are commonly illustrated with discrete, countable spatial expanses—formalizes the relation 'is a *piece* of' rather than the common sense part relation 'is part of.'

classes according to homomerity pattern, participants structure, dynamic composition, dynamic shape, and dynamic context. Within this rich typological structure of free processes one can identify interactional structures that are suitable counterparts for many common sense genera nouns, not merely the traditional explananda of 'thing,' 'property,' 'state of affairs,' or 'person,' but also the species of disposition: capabilities, capacities, tendencies, or propensities.[59]

APT has a number of explanatory assets—it is well-founded, formally simple, and monocategorial. Due to the features of its basic category, APT affords new and, as I have argued elsewhere, successful strategies in addressing the core problems of the ontological tradition (the status of universals, persistence, thinghood, predication etc.). Surprisingly, our judgments about the numerical, qualitative, and transtemporal identity of things and persons can best be captured if we treat them as statements about types of dynamics or free processes. Even though the primary domain of application of APT is the ontological interpretation of common sense reasoning as embedded in a (Indo-European) natural language, I believe there are a variety of aspects of the scheme that could make it attractive to those in search of an ontology for QFT.[60]

Of crucial importance in this regard is the APT-distinction between an *amount* and a *quantum** of a free process, extrapolating from observations for common sense stuffs (3-dimensional non-countables) to free processes in general. Stuffs occur in amounts with determinate spatio-temporal location: in my cup there is an amount of coffee. The amount of coffee in my cup is a fairly determinately bounded and spatio-temporally located particular: this amount of coffee does not exist anywhere else at the same time. Similarly, consider Bach's *Magnificat* resounding now in my living room or the snowing I see through my window while listening. These amounts

[59]For details cf. Seibt 1996, 2000c, 2001, and in particular 2002.

[60]I should also mention one feature that might make the scheme right away far less attractive than Whitehead's. The present version of APT assumes spacetime as a 'given'. Based on a modified condition of homomerity which does not make reference to spatial parts, I do categorize spacetime as the unique free process that is ultimately homomerous (every part is of the same kind as the whole); also, free processes are 'subjectless' also in the sense that they are not modifications of spacetime but just 'forms of dynamics.' But at the moment I doubt that it is possible (without adopting a theory of circular definitions as set out in Gupta and Belnap 1993) in APT to define spacetime points in terms of differences in physical content, i. e., in terms of different interaction 'products' of free processes (cf. option (b2) in section 3.4 above).

of classical music and snowing are particulars: the piece playing in your living room, the snowing in front of your window are different amounts of classical music or snowing. In general, amounts of free processes are particulars—and amounts are the only type of particular there is in APT.

It is crucial to see, however, that amounts, though particulars, are not necessarily fully determinate: in my cup is a (determinable) amount of liquid, in my living room is a (determinable) amount of music. Most amounts of stuffs and other non-countables are—*pace* substance-ontological principles CAP-9 and CAP-10 above—*concrete particular determinables*. Finally, note that amounts are ordinal-countable in the sense defined in section 3.3 above. Since amounts have definite spatiotemporal locations it does make sense to speak about the first cup of coffee, the second replay of the recording and the third amount of accompanying snowing. In sum:

> (P4) Amounts of free processes are more or less determinate, ordinal-countable particulars.[61]

In the discussion of the ontology of stuff terms the notions of 'amount' and 'quantity' are used interchangeably to refer to particular portions of a stuff.[62] But in this way one is likely to lose sight of an important distinction which I shall couch here in terms of a contrast between the amount of coffee in my cup and the 'quantum*' ('as much') of coffee in my cup. If we measure an amount of stuff S by determining how many measuring units of measurable property such as volume, weight, or temperature are contained in the amount, we are interacting with quanta* of S. The primary sense of 'as much' may be tied to volume but we can generalize the notion and speak of a d-quantum* with d ranging over the set of measurable properties d_i. Assume that my cup of coffee, an amount of coffee '$[\alpha]$', is a quantum* of volume with the value of 2 dl, a quantum* of weight with the value of 200 g, a quantum* of temperature with the value of 50^0C, and that volume, weight, and temperature are the only measurable properties of $[\alpha]$. The ontological assay of $[\alpha]$ could then be written as an ordered tuple of relevant d-quanta* q_i with a spacetime location r, i.e., '$[\alpha] = \langle q_1, q_2, q_3, r\rangle$.' A quantum* is

[61]An amount α of a free process β occurrent in spacetime region R is in APT defined as the γ which is the interaction of β with a free process ρ_i which is (or; represents) region R (cf. previous footnote).

[62]The term 'quantity' is thus used rather differently by ontologists and physicists.

not a value of a measurable property but 'something which gives rise to or affords that value in a measuring process.'[63]

Quanta* are concrete yet indeterminate in the following three respects. First, quanta* do not have the kind of spatio-temporal location characteristic of particulars. They are not at precisely one spatial location at any moment in time; they are not, as I shall say, 'uniquely located'. Consider q_1, the quantum* of coffee-qua-volume in my cup, whose value is 2 dl. Could we say that q_1 is composed of two component quanta* q_{1a} and q_{1b} whose value is 1 dl? As an assertion about the spatial composition of q_1 such a statement would not make any sense. In saying that q_1 has the value 2 dl we do not thereby imply that there are two uniquely localizable deciliters sitting neatly next to each other. We are not even saying that there are two uniquely localizable deciliters that might "switch places without empiral consequences." In saying that q_1 has the value 2 dl we are rather saying that a quantum $q*$ of value 1 dl is occurring twice in $[\alpha]$. This is an assertion about relationships between values of measurements but it is also an assertion about that which gives rise to or affords such measurements, which is concrete but non uniquely localized and thus not particular.

Second, since quanta* are not uniquely localized it follows that they are not ordinal-countable but merely measurable with continuous or discrete spectrum of values (cardinal-countable).

Third, a quantum* of stuff S is more indeterminate than an amount of S since it is a projection of S onto a certain evaluative dimension, e. g., coffee-qua-volume, coffee-qua-weight, coffee-qua-temperature. There are features of S that do not enter into that which affords a certain measurement value for a measurable property of S: not all of that which gives rise to the measurement of the color of coffee is that which gives rise to the measurement of the volume.

Finally, the perhaps most important aspect of indeterminateness in quanta* of classical stuffs is that they are what 'gives rise to or affords' certain measurement results. We may be more used to think only of 'secondary qualities' as 'dispositional'—that which 'gives rise to a certain measurement value of' sweetness is not itself sweet—but there is no good reason to exclude 'primary qualities.' In APT all quanta* are characterized as capacities. The various idioms of capacity (give rise to, afford, engender, to

[63]For the notion of 'affordance' and its possible application for the interpretation of QFT cf. Harré 1988.

be potentially etc.) can be painlessly analyzed within a process-ontological setting. They are statements about an antecedent functional stage s_i of a process β characterized, with prospective focus, in terms of a consequent functional stage s_j of β relative to a certain dynamic context. In other words, when we speak of capacities or potentials we speak of a functional stage of a process type in terms of its contextualized continuation. As that which gives rise to measurement values of measurable properties of S, a quantum* of S is merely an indeterminate stage in a measurement process yielding a certain value of a measurable property of an amount of S. To sum up:

> (P5) A quantum* of S is an individual entity but not particular and not ordinal-countable; it is an indeterminate functional stage in a process (type) β which is the measurement of a value for a measurable property of S.

These are sketchy remarks but I hope the following has become clear. APT countenances individuals that are non-particulars (quanta*) and individuals that are particulars and ordinal-countable (amounts). Both quantities and amounts lie on a gradient of (in)determinateness. Traditional particulars as assumed within the substance paradigm (i. e., the fully determinate, individual, and ordinal-countable particulars that are allegedly the ontological counterparts of 'classical particles') exist in APT only as limit cases, as ultimately determinate amounts of free processes.

On the basis of (P4) and (P5) let me try then and bring into view some conceptual affinities between APT and the Fock space formalism of QFT. In APT a complex free process (e. g., a blizzard, a fugue, the stock market) is the interaction process of component processes, e. g., the inter-activity* α of component activities* β_i: '$\alpha = \mathrm{In}\langle\beta_i, \beta_2, \ldots \beta_k\rangle$.'[64] As a dynamic 'mixture' of dynamic 'stuffs' a complex process might in particular be a superposition of activities* (i. e. dynamics which can be represented as the harmonic modes of a classical field). Assume then, as above, that an amount of α is described in terms of a list d-quanta* of the β_i, $[\alpha] = \langle q_1(\beta_1), q_2(\beta_2), \ldots q_k(\beta_k)\rangle$, and that such quanta* have discrete values. For any measurable property A of the β_i we can then determine a distribution patterns of A within α representing A-quanta* of the β_i in

[64]The term 'activity*' refers to non-telic free processes akin to common sense activities, in contrast with telic occurrences: productions, developments, events etc. All simple free processes are activities*; complex free processes may or may not be activities*.

terms of their discrete values: $F_A = \langle n_1, n_2, \ldots n_k \rangle_A$. Under the given assumption a complex process α thus can also be represented in terms of a list of distribution pattern F_d for any measurable property d of the component activities* of α.

Drawing the analogy along a different route, to accommodate time-indexed states, let $[\alpha]$, i. e., the amount of a complex process α be analyzed in terms of component amounts $[\beta_k *]$ with minimal temporal extent. For example, if $[\alpha]$ is a particular performance of a fugue with three voices, we can take $[\alpha]$ to consist momentary complex sounds $[\beta_k *]$ which each consist of 3 momentary single sounds $[\gamma_m]$ representable by three quanta* of the measurable properties (observable) pitch, timbre, and intensity: $[\gamma_m *] = \langle q_p, q_t, q_i \rangle$. We denote again quantities by their values: $[\gamma_m *] = \langle n_p, n_t, n_i \rangle$; *vice versa*, for each observable A (pitch, timbre, or intensity), we can state how A is distributed over the single sounds: $\langle n_1, n_2, n_3 \rangle_A$, where n_i is the value of the quantity q_A in $[\gamma_i *]$.

These are two routes, then, to formally represent complex free processes. Despite obvious disanalogies (ordered lists instead of vectors, discrete classical values instead of eigenvalues of an operator, etc.) the analogies are, I hope, sufficiently promising to explore an APT-interpretation of QFT. If we interpret field-quanta as APT-quanta* we avoid two complications arising for Teller's quanta. Teller's quanta are said to be particulars; thus, unless we are given an alternative definition of particularity, they must be uniquely located. But even if Teller's quanta have a "high degree of localizability" by occurring in "well-defined spacetime volumes"[65] there is no way to ensure the required uniqueness of spatial occurrence—individuated by descriptive thisness only, a Tellerian quantum can occur in several spatial regions at once. APT-quanta* on the other hand are not particulars. They are not uniquely but only generally located. Further, Teller suggests that superpositons of exact number states, that is, states with indefinite numbers of quanta, can be interpreted as "propensities to reveal one of the superimposed properties under the right 'measurement' conditions."[66] More precisely, "two descriptions each characterizing (certain or merely probable) manifestation of quanta can superimpose to characterize a state in which either of the two manifestations will occur with probabilities calcu-

[65]Teller 1995: 106.

[66]Teller 1995: 32.

lated from the superposition formalism."[67] This move is suspect since the propensities in question concern the very existence of the entities (quanta) to which the propensity is ultimately ascribed, given that states are collections of quanta.[68] More importantly, it seems to go beyond the empirical content of QFT:

> The occurrence of quanta, by which a superposition state is realized, are single measurement events, and the theory—according to our current knowledge—is not able to account for single measurement events. The ontology for QFT should not be built upon a type of entity the presence or absense of which is not a consequence of the theory.[69]

In APT on the other hand dispositions or conditional dynamic structures figure already prominantly in the description of classical domains. A quantum* of β is the indeterminate antecedent stage of a specific measurement process for a measurable property on an amount of β. In APT a quantum* of β can have indefinite value ('n or m'); in this case it is the indeterminate antecedent stage of an inspecific measurement process (e. g., characterized as the disjunctive process of *measuring for* d_1 *or measuring for* d_2, or as the disjunctive process of *measuring* n *for* d_1 *or measuring* m *for* d_1). If the quanta* of β have definite values, β is a more specific process; if they have indefinite values, β is a less specific process (not a fugue but merely music). From the point of view of APT superpositions of states with exact numbers of quanta are to be taken as dynamic states or activities* with an indefinite number of quanta*, i. e., as general or indeterminate dynamic states. In general, once the traditional substance-ontological link between individuality and determinateness or specificity is eliminated, the superposition of probable specific or determinate states can be read as denoting a general or inspecific state. Finally, by admitting individuals that are indeterminate dynamic conditions (e. g., processes with disjunctive continuations), APT offers a new way to make ontological sense of probability statements without invoking dependent entities such as propensities which require a 'bearer' or subject. (In this way the vacuum considered as a state with zero

[67] Ibid. 105f.

[68] Cf. Bartels 2000: 335

[69] Ibid.

quanta yet positive averages (expectation values) of field quantities is no longer a puzzle of unborn propensities.)

3.6 Conclusion

To restate the *caveat* of the introduction, the primary aim of the considerations presented here is to assist rather than to present research on the ontological interpretation of QFT. I have drawn attention to the fact that contemporary ontology is heavily influenced by dispensable presuppositions of the substance paradigm. In consequence, ontological research, whether on classical domains or others, has explored only a small subsection of the conceptual space available for category definition. The debate about the ontological interpretation of quantum domains has so far rather conservatively adopted the ideas and conceptual options developed within the ontological research tradition (such as the incoherent notion of 'primitive thisness'). Instead of importing substance-ontological presuppositions the debate about the ontological interpretation of quantum domains should pursue new constructional paths. In fact, the ontological interpretation of quantum field theoretic notions already has begun with the revision of substance-ontological principles. Paul Teller's suggestion that quanta are particulars yet do not posssess primitive thisness amounts to an attempt to sever the traditional link between particularity and individuality. It is important to note, however, that Teller's suggestion is not yet embedded in a full-blown ontological theory; thus strictly speaking Teller's interpretation of QFT introduces a new type of entity but cannot yet count as an ontological interpretation of QFT. There are, however, at least three systematic frameworks that might prove suitable candidate-ontologies for QFT. In keeping with the particularist tradition in ontology, one might pursue, first, a trope-theoretic reformulation of Auyang's event ontology or, second, a Whiteheadian theory of occasions. But the most promising route for philosophers of physics might be to develop ontological systems that even further deviate from substance-ontological habituations and even abandon the particularist stance still endorsed in trope theory or the theory of occasions. The theory of free processes sketched in the last section might serve as an example for the degree of 'conceptual liberation' that is possible within the ontology for classical domains, and perhaps necessary for the ontology of QFT. One crucial requirement of ontological research, however—the only one, incidentally, which can account for the popularity of the substance paradigm—must not be overlooked in exploring the uncharted regions of category space. Ontological categories have to be well-founded. The task of ontology is not only to describe the domain of a theory, but to offer an explanatory description whose basic terms we can 'agentively understand'.

References

Albert, D. and B. Loewer (1988): "Interpreting the Many-Worlds Interpretation," *Synthese*, 77: 195–213.

Bacon, J. (1995): *Universals and Property Instances: The Alphabet of Being*, Oxford: Blackwell.

Baker, R. (1967): "Particulars: Bare, Naked and Nude," *Nous*, 1: 211–212.

Barnette, R. (1978): "Does Quantum Mechanics Disprove the Principle of the Identity of Indiscernibles?," *Philosophy of Science*, 45: 466-470.

Bartels, A. (2000): "Quantum Field Theory: A Case for Event Ontologies?" in: Faye, Scheffler and Urchs 2000, 327–343.

Bergmann, G. (1967): *Realism*, Madison: University of Wisconsin Press.

Brown, H., and R. Harré (eds.) (1988): *Philosophical Foundations of Quantum Field Theory*, Oxford: Clarendon.

Cao, T. Y. (ed.) (1999) *The Conceptual Foundations of Quantum Field Theory*, Cambridge: Cambridge Unversity Press.

Faye, J., Scheffler, U., and M. Urchs (eds.) (2000): *Things, Fact, and Events. Poznan Studies in the Philosophy of the Sciences and the Humanities*, vol. 76, Amsterdam: Rodopi.

French, S., and D. Krause (1999): "The logic of quanta," in: Cao 1999, 324–341.

French, S., and M. Redhead (1988) "Quantum physics and the idenitity of indiscernibles," *British Journal for the Philosophy of Science*, 39: 233–246.

Fine, K. (1991): "The Study of Ontology," *Nous*, 25: 263–294.

Folse, H. (1982): "Complementarity, Bell's Theorem, and Process," *Process Studies*, 12: 259–273.

Gupta, A., and N. Belnap (1993): *A Revision Theory of Truth*, Cambridge: MIT Press.

Hättich, F. (2000): "A Whiteheadian Interpretation of Algebraic Field Theory," unpublished manuscript of a Ph. D. dissertation at the University of Bonn.

Harré, R. (1988): "Parsing the Amplitudes," in: Brown and Harré 1988, 59–73.

Haslanger, S. (1989): "Persistence, Change, and Explanation," *Philosophical Studies*, 56: 1–28.

Hochberg, H. (1992): "Troubles With Tropes," *Philosophical Studies*, 59: 18 f.

Kuhlmann, M. (2000): "Processes as Objects of Quantum Field Theory: Consequences for the Interpretation of QFT," in Faye, Scheffler and Urchs 2000, 365–389.

Kuhlmann, M. (2001): "Quanta and Tropes. Trope Ontology as Descriptive Metaphysics for Quantum Field Theory," in: Meixner, U. (ed.), *Metaphysik im post-metaphysischen Zeitalter/ Metaphysics in the Post-Metaphysical Age*, Wien: Öbv & hpt.

Lango, J. (1972): *Whitehead's Ontology*, Albany: SUNY.

Mormann, T. (1995): "Trope Sheaves. A Topological Ontology of Tropes," *Logic and Logical Philosophy*, 3: 129–150.

Puntel, L. (1993): "The Context Principle, Universals and Primary States of Affairs," *American Philosophical Quarterly*, 30: 123–135.

Redhead, M. (1988): "A Philosopher Looks at Quantum Field Theory," in Brown and Harré 1988, 9–23.

Rovelli, C. (1999): "'Localization' in quantum field theory: how much of QFT is compatible with what we know about space-time?" in Cao 1999, 234–245.

Schurz, G. (1995) "Ein quantenmechanisches Argument für die Existenz konkreter Universalien," in Brandl, J., Hieke, A., and P. Simons (eds.), *Metaphysik. Neue Zugänge zu alten Fragen*, Volume 11 of *Conceptus Studien*, St. Augustin, Academia Verlag.

Seibt, J. (1990a): *Properties as Processes. A Synpotic Study in Wilfrid Sellars' Nominalism*, Reseda: Ridgeview Pub. Co.

Seibt, J. (1990b): *Towards Process Ontology: A Critical Study in Substance-Ontological Premises.* Michigan: UMI-Dissertation Publications.

Seibt, J. (1996a): "Individuen als Prozesse: Zur prozess-ontologischen Revision des Substanzparadigmas," *Logos*, 5: 303–343.

Seibt J. (1996b): "Non-countable Individuals: Why One and the Same Is Not One and the Same," *Southwest Philosophy Review*, 12: 225–237.

Seibt, J. (1996c): "The Myth of Substance and the Fallacy of Misplaced Concreteness," *Acta Analytica*, 15.

Seibt, J. (1996d): "Existence in Time: From Substance to Process," in: *Perspectives on Time*. Boston Studies in Philosophy of Science, ed. by Faye, J., Scheffler, U., and M. Urchs, Dordrecht: Kluwer, 143–182.

Seibt, J. (1999): "Dinge als Prozesse," in: Hüntelmann, R., and E. Tegtmeier (eds.), *Neue Ontologie und Neue Metaphysik*, Köln: Academia Verlag, 1999, 11–41.

Seibt, J. (2000a): "Constitution Theory and Metaphysical Neutrality: A Lesson for Ontology?" *The Monist*, 83: 161–183.

Seibt, J. (2000b): "Ontology as Theory of Categorial Inference," in: Peres, C., and D. Greimann, (eds.), *Wahrheit-Sein-Struktur. Auseinandersetzungen mit Metaphysik*, Hildesheim: Olms, 272–297.

Seibt, J. (2000c): "The Dynamic Constitution of Things," in: Faye, Scheffler and Urchs 2000, 241–279.

Seibt, J. (2001): "Formal Process Ontology," in: Welty, C. and Smith, B. (eds.), *Formal Ontology in Information Systems. Collected Papers from the Second International Conference* Ogunquit: ACM Press, 333-345. ACM Digital Library.

Seibt, J. (2002): *Free Processes: A Study in Ontological Revision*. Habilitationsschrift submitted at the University of Konstanz.

Sellars, W. (1963): *Science, Perception, and Reality*, London: Macmillan.

Shimony, A. (1965): "Quantum Physics and the Philosophy of Whitehead," *Boston Studies in the Philosophy of Science*, vol. 2, ed. by Cohen, R., and M. Wartofsky, New York: Humanities, 307–330.

Simons, P. (1994): "Particulars in Particular Clothing: Three Trope Theories of Substance", *Philosophy and Phenomenological Research*, 54: 553–575.

Stapp, H. (1977a): "Theory of Reality," *Foundations of Physics*, 7: 313–323.

Stapp, H. (1977b): "Quantum Mechanics, Local Causality, and Process Philosophy," *Process Studies*, 7: 173–182.

Stapp, H. (1979): "Whiteheadian Approach to Quantum Theory and the Generalized Bell's Theorem," *Foundations of Physics*, 9:1–25.

Stegmaier, W. (1977): *Substanz: Grundbegriff der Metaphysik.* Stuttgart: Fromann-Holzboog.

Stöckler, M. (1988): "Individualität, Identität, Ununterscheidbarkeit: Überlegungen zum Gegenstandsbegriff in der Quantenfeldtheorie," *Conceptus*, 22:1–26.

Teller, P. (1990): "What the Quantum Field Is Not," *Philosophical Topics*, 18: 175–187.

Teller, P. (1995): *An Interpretive Introduction to Quantum Field Theory.* Princeton: Princeton University Press.

Weinberg, S. (1999): "What is QFT and what did we think it was?" in Cao 1999, 241–251.

Whitehead, A. N. (1929): *Process and Reality*, New York: Macmillan.

Zemach, E. (1970): "Four Ontologies," *Journal of Philosophy*, 23: 231–247.

This page is intentionally left blank

Chapter 4

Analytical Ontologists in Action: A Comment on Seibt and Simons

Meinard Kuhlmann
University of Bremen

Abstract. *Since I like (trope) bundle theories, everything in my comment will come in bundles, namely in bundles of three. I will have three questions, three points in answering each of these questions and three arguments against each of Johanna Seibt and Peter Simons. My first question is 'Where do Seibt and Simons agree or at least converge?', and the second one is 'Where do Seibt and Simons disagree or at least proceed differently?' (section 4.1). The third question is 'Where do I disagree with Seibt and Simons?' (sections 4.2 and 4.3). I will conclude my comment with an evaluation of the significance of the papers by Seibt and Simons for the philosopher of science who is concerned with the ontology of quantum field theory (QFT).*[1]

4.1 Ontology and Physics: Seibt Versus Simons

Regarding methodology, as well as regarding some fundamental evaluations, there are various issues on which Seibt and Simons agree, or at least converge. As analytical ontologists, both Seibt and Simons stress the formal (or analytical) component of ontology, on the one hand, and the significance of looking at special sciences on the other.[2] The emphasis on empirical issues

[1]Since Simons refers frequently to his own previous publications I will proceed correspondingly by using these publications as background information for my comment. In particular, I have been using Simons 1994, 1998a and 1998b. Seibt's paper in this volume, in contrast to Simons', is far more self-contained, which explains the considerable difference in the lengths of both papers. Nevertheless, in Seibt's case I have been writing with an eye on her other publications as well, in particular Seibt 1995, 1996 and 2001.

[2]In short, analytical ontology can be described as the rehabilitation of some old metaphysical questions which are treated in an analytically purified way. 'Analytically' here

is of particular importance for a volume on the ontology of QFT since it renders analytical ontologists particularly capable for a collaboration with physicists and philosophers of physics. Seibt points out that ontological examinations are inherently relative, namely relative to a given "data set of categorial inferences", be it in natural languages or in formalisms like the ones used in Quantum Field Theory (QFT). 'Categorial' refers to categories such as properties and substances (or things). Seibt supplies some simple examples in her footnote 4. The task of the ontologist, according to Seibt, is to explain the validity of the categorial inferences by supplying an appropriate ontology.

Simons, elaborating on ideas from Christian Wolff, Husserl, Whitehead and D. C. Williams, proposes to distinguish two parts *within* metaphysics, namely (formal) ontology and (metaphysical) systematics.[3] According to Simons' view ontology is the formal part of metaphysics, concerned with the establishment of an array of categories that can be applied to all there is. (Metaphysical) systematics, on the other hand, is that part of metaphysics which deals with the contents to which the formal categories are to be applied. Simons describes a number of tasks for (metaphysical) systematics. It has to check whether the categories supplied by ontology are applicable to the diversity of being, whether the categories are appropriate and whether they are complete in order to catch everything we encounter in the world. In other places[4] Simons has called the last task the 'integration requirement'. Simons' integration requirement is crucial in the present context since it explains his interest in physics as well as his reservation about putting to much emphasis on physics. It accounts for Simons' interest in physics simply because the objects which are studied by physics are an important part of the world, and the discoveries and theories of physics belong to the most refined and best-checked findings of mankind. Nevertheless, Simons

refers to the use of modern logic (which emerged at the end of the nineteenth century) for the analysis of language. In addition to its methodological aspects, it stands in the analytical tradition by its high esteem for exact empirical sciences. However, in distinction from the original attitude in logical positivism, analytical ontologists hold that the emphasis on empirical scientific results leaves room for a contribution by philosophy which is not purely methodological. For a brief account of that contribution see the Introduction to this volume.

[3]Simons has borrowed the term 'systematics' from biology, where it is used to classify the living world.

[4]See, e. g., Simons 1998b, section 3.

limits the attention metaphysics should devote to physics since, according to his integration requirement, metaphysics has to account for all there is in the world, not only for physics and its objects. I will come back to this point in section 4.3 as well as in my conclusion with a critical evaluation.

While Seibt and Simons generally agree on the importance of empirical data for the ontologist, they lay emphasis on different roles. Whereas Seibt stresses that ontological investigations should start with observations (about the set of categorial inferences licensed by a given language or theory), Simons puts the establishment of ontological systems first and uses observations and empirical results as tests for the ontologies.

The second issue I will briefly deal with is a comparison of how Seibt and Simons proceed in finding an appropriate ontological framework for QFT. As is quite typical for analytical ontologists, both authors consider a list of possible candidates and try to show that all alternatives but one can be excluded—although they come to different conclusions as to which the remaining one is. Seibt as well as Simons use two types of arguments in their evaluation of candidate ontologies. The first type of argument refers to the internal consistency of ontological theories. The second kind of argument leaves the range of internal arguments and looks at how well different ontologies do with respect to the everyday world and the sciences.

Nevertheless, there are some differences in how Seibt and Simons find and discard possible ontologies. While in this respect Simons does not seem to follow a definite procedure, Seibt proceeds in a way she calls "axiom variation". A large proportion of her work consists in considering different combinations of category features. One result is that the most natural ways to conceive of the notion of substance lead to inconsistencies, e. g., the combination of persistence, subjecthood and determinateness. Seibt sets out that some proposals for an ontology of QFT might be excluded in a corresponding way. If Teller's conception of quanta[5] rests on the assumption of particularity, discreteness and a certain sense of non-countability, then, Seibt argues, one gets an inconsistency just by combining these three category features.

The third and last point of comparison deals with the final conclusions both authors arrive at, a critical evaluation of which will follow in the next two sections. While some ontological frameworks are more problematic for internal reasons and others more for the lack of 'data fitting', Seibt as well

[5]Teller 1995, Ch. 2.

as Simons come to the conclusion that for both kinds of reasons substance ontologies have no chance of surviving the competition of ontologies. The "myth of substance" (Seibt) rests on the "unacceptably anthropocentric" thesis of a "pre-established harmony of linguistic and ontological categories" (Simons). And this is the third and last similarity of Seibt and Simons I wish to highlight. Leaving internal issues of inconsistencies of substance ontologies—e. g. due to their overdetermination—aside, both authors stress that quantum physics yields strong arguments against taking the notion of substance as the basic category of one's ontological framework. The impossibility of conceiving of 'identical particles' in a compound state as individuals speaks against the applicability of the notion of substance since quantum particles should be the prime cases of substances in the realm of quantum physics.[6] While Seibt and Simons have various similarities in their negative results, the most important divergence between Seibt and Simons is that the final (positive) results of their investigations are different. In the end they favour different ontologies, although both can be classified as 'revisionary' (in contrast to 'descriptive') in the sense that they diverge from conceptions which (seem to) give an immediate ontological description of, e. g., the categorial structure (properties, things ...) of our language. Seibt and Simons argue in favour of their proposals by considering QFT along with other more general considerations. Seibt defends a certain kind of process ontology, which she calls "Axiomatic Process Theory (APT)", and Simons holds an 'ontology of invariant factors'. In the next two sections I will give a critical discussion of these theories in turn.

4.2 Processes and Ontological Parsimony

Seibt describes her "Axiomatic Process Theory (APT)", or alternatively her "theory of free processes", as the strongest rejection of traditional substance-ontological schemes, ingredients of which she sees still at work in most if not all other approaches. The last, rarely questioned presupposition of substance ontologies one should abandon, according to Seibt, is to assume that being a concrete individual would imply particularity. While a sufficient condition for individuality is to be reidentifiable, par-

[6]I leave aside the interesting question of whether there are other, more appropriate candidates for substances in quantum physics. For further alternatives see, e. g., Scheibe 1991.

ticulars are contrasted with universals (which can be instantiated or realized many times). Seibt motivates her approach by the observation that there are various examples of individuals that are not particulars, like a wedding or snowing. "They are individuated in terms of their descriptive thisness, not by spacetime location, and may occur in a multiply disconnected spatiotemporal region with fuzzy boundaries."[7] Extrapolating from these observations, Seibt proposes to take as the basic entities of one's ontology free processes in the sense of individuals which are concrete (i. e. spatiotemporally occurent) but which are not particulars.

In my view, the most important deficiency of a process-ontological approach to QFT is the lack of a satisfactory explicit description and definition of the assumed basic processes. As has been stressed various times, the "processes" depicted in Feynman diagrams cannot be understood in a realistic way which would make them candidates of basic processes. For mathematically-minded physicists there is the question of a mathematical definition and a concise description of the mathematical structure of the set of processes. A first idea would be to understand a process as the triple of two events and a unitary time evolution operator. A good starting point could be to explore whether and where conventional conceptions of processes differ from the kind of processes which a process ontology postulates. An interesting subquestion to the first one is the connection of process ontology to recent theories of the structure of spacetime (e. g. geometro-dynamical models).[8]

The second reservation against Seibt's process ontology is concerned with explanations for phenomena that are natural for a substance ontologist, while they call for a lot of effort on the part of the process ontologist. Whereas the substance ontologist has a hard time to explain how change in time is possible, even though the things which change supposedly keep their identity, the process ontologist has the opposite problem: Why do we have the strong impression that many things are more or less static if everything is composed of processes? Why does it appear that stable particles and molecules exist? One possible explanation is to assume the existence of counterprocesses that exactly balance other processes, with the overall

[7]*See section 3.5 in the present volume.*

[8]David Finkelstein made some interesting proposals for a process-ontological interpretation (or better: revision) of quantum physics. For details see Finkelstein 1996 as well as various of his articles from the last thirty years cited therein.

effect of the appearance that nothing happens.

This brings me to the third and last objection to Seibt. Leaving questions of internal consistency aside, I doubt that a process ontology rates very well with respect to ontological parsimony. If the above process-ontological account of more or less static objects is correct then one would have to assume a plethora of processes, that are not even observable as processes. in which process ontology expends more than necessary. Maybe one could call it revisionary parsimony.

4.3 Tropes, Invariant Factors and Quantum Field Theory

One kind of 'candidate general ontologies for situating quantum field theory' that Simons discusses are trope ontologies. Simons himself has gained a certain fame for the elaboration of this type of ontology, in particular with his much-discussed paper "Particulars in Particular Clothing: Three Trope Theories of Substance" (1994). In his contribution to the present volume, trope ontology appears just as one among five other ontological frameworks—which are all dismissed in favour of a sixth possibility, a so-called "ontology of invariant factors." Since I myself am sympathetic to trope ontology, partly because of Simons' very persuasive 1994 paper, I was somewhat surprised to find him arguing mostly against a tropes-only ontology now. The main reason why I am surprised is that I cannot find compelling reasons for Simons' seeming deviation from his previous point of view, as far as I had understood it. For my taste Simons has now adopted a weaker position than he already had. Apart from philosophical aspects, I think that Simons' previous position (as I took it) allowed for a very smooth and illuminating link-up with fundamental theories in physics, while his new point of view comes at the expense of much explanatory power, at least as far as ontological considerations with respect to QFT are concerned. The three arguments I wish to present against Simons are, first, a philosophical argument against Simons' argument against a "tropes-only ontology"; second, an argument in favour of a way Simons rejects; and, third, an argument concerning the fitness of Simons' ontology for giving an ontological account of QFT.

After discussing various advantages of trope ontologies Simons ends section 2.6 of his paper in this volume with the comment that "a tropes-only ontology is insufficient for several reasons", a claim that Simons promises to meet in section 2.8 when introducing and defending his current proposal of an ontology of invariant factors. The only proper argument against a 'tropes-only ontology' I can find in section 2.8 leans on Bolzano and runs as follows:

> If something is dependent, then [...] there must be at least one thing that is independent, even if it is the whole world. [...] So there cannot be a tropes-only ontology: the world is not a trope, and the world is not nothing.[9]

I find this argument weak[10] since it hinges on a premise which is controversial in itself and which Simons does not even mention in this context. The implicit premise I have in mind is the impossibility of ontological reduction.[11] I think that holding a 'tropes-only ontology' is by no means tantamount to claiming that all there is in the world are tropes. I understand it that the claim of a 'tropes-only ontologist' is merely that everything can be *reduced* to tropes which are the fundamental entities in the world. For this reason, I take it, Campbell calls it a 'one-category ontology'[12] thereby stressing the intended ontological parsimony of the approach. All other non-fundamental entities are analyzed *in terms of* tropes and of course don't need to be tropes themselves. In my view, this is the very aspect that makes trope ontology an ontological theory with explanatory power. The fact that there are things in the world which are not tropes does not

[9]See section 2.8 in the present volume.

[10]Besides my main concern I wonder whether the argument is sound at all—or at least whether it is complete. Consider an uncle and his nephew in a room and assume a substance ontology according to which human beings are irreducible entities. Would we say that there are *three* things in that room, the uncle, the nephew *and* the uncle-nephew-couple? If one were willing to concede that one would be forced to say that a 'human-beings-only ontology' is inapproriate for the contents of that room because a couple is not a human being and a couple is not nothing. But maybe Simons would say so. I am sure that Simons can cope with my objection but as his text stands I don't find it convincing in this respect.

[11]Simons only makes a passing reference to his anti-reductionist attitude in the "Methodological Preamble" at the beginning of his paper. In other publications Simons explictly places himself in the phenomenological tradition of Brentano and Husserl and he defends an anti-reductionist position as phenomenologists typically do. The phenomenologist wants to retain the diversity of appearances (or phenomena) instead of trying to reduce them to a set of basic entities.

[12]Ch. 1 in Campbell 1990 provides an extensive exposition and defense of trope ontology, which is partly the reason for the new interest in trope ontology in the last 10-15 years.

speak against but in favour of trope ontology, provided that trope ontologists can tell a convincing story in terms of tropes. This is how I understood Simons "nuclear theory" of tropes.[13]

If I have overinterpreted Simons 1994, as I suppose now, this seems a good place and time to point to one advantage that a 'nuclear theory of tropes' has, or better could have, seemingly only if it is not understood in Simons' way. This is my second argument against Simons' rejection of a tropes-only ontology. I think that such a theory, based on quantum physics, can deal with the so-called 'boundary problem' in a very convincing way.[14] The boundary problem—which holds potential danger for trope ontologies—has to do with the delimitation of tropes. Suppose you have a piece of blue paper and you try to give an account of it in terms of the classical bundle theory of tropes. It suggests itself to say that this piece of paper is the bundle of this particular blue trope, this particular consistency trope, and so on for all other properties. Now one may ask what happens, e. g., to the blue trope if the paper were cut into two pieces. Since both pieces have their own blue tropes now, one gets an inconsistency with the trope-ontological account for the initial piece of paper. Why is the blue of, say, the left half of the paper considered to be an entity of its own only after the paper is cut into two pieces although it has not at all changed by the cutting. The problem is to determine the boundary where one trope ends and the next one begins.

I can see two possibilities to handle the boundary problem. One is advocated by Campbell (1990, sec. 6.8), which he calls the field approach. According to this approach, the only true tropes are fields which are spread out through the whole universe. The boundary problem is solved since anything that is less extended than the whole universe cannot be a proper trope. I think building on Simons' 'nuclear theory of tropes' one could have a second solution to the boundary problem. According to this approach, only tropes in bundles corresponding to elementary quantum objects are fundamental tropes, e. g. charge tropes or spin tropes. All other tropes,

[13]See Simons 1994, pp. 567-574. In the present paper Simons summarizes his 'nuclear theory of tropes' (without mentioning his previous term for it) in two sentences by saying that "a concrete individual [is a] complex of tropes [which consists] of an inner core of tightly co-dependent tropes constituting the individual's "essence" and a corona of swappable or variable adherent tropes allowing it to vary its intrinsic features while remaining in existence."

[14]Cf. Campbell 1990, Ch. 6.

like the green trope of this cup, are merely good for illustration of how to understand trope ontology. Whether the fundamental tropes are field-like or not is another question, one that is directly connected with various discussions, e. g., in this volume. In any case, one is no longer troubled by the boundary problem because these fundamental tropes cannot be cut into parts any more. This is how I think a (let us call it) 'nuclear tropes-only ontology' could be started.

Finally, I wish to raise an objection against the ontology that Simons eventually favours, namely the "ontology of invariant factors". Without going into too much detail, I think one example is very telling. According to Simons, the modal configuration of an electron is as follows.[15] In general, the modes of an electron are "thetic, aphairetic, in different phases of the life-cycle passing through all 3 bias modes, absolute, hæccic, unicate, bracteal, adiaphoric." The other modes depend on the circumstances. If it is an isolated electron it (i. e. its modal dimension 'valence') is plene, if it is an entangled electron it is moietic. An "electron qua wave [is] heteronomous (disturbance in a field) or autonomous (a self-sufficient process) depending on [the] theory[, an] electron qua particle [is] autonomous (a substance) or heteronomous (an invariant across a process) depending on [the] theory". For an explanation of these terms the reader may wish to consult Simons' contribution to this volume as well as previous papers (in particular Simons 1999) as cited therein.

Simons' example reminds me of the situation in high energy physics in the 1950s and 1960s when the 'particle zoo' contained some 500 'elementary particles' which resulted in a widespread feeling that one had not reached the bottom yet. In Simons' categorial scheme there are even 3,072 (!) fundamental modal combinations. Of course this is not a knock-down argument, but I think it is quite a negative feature if the set of elementary entities becomes too cumbersome. This is even more so if there are more convincing competitors around and I think trope ontology—as briefly depicted above—is one of them.

Treating baroque music and quantum field theory ontologically on a par, as Simons tries to do, is certainly elegant in the sense that one needs just one categorial scheme. However, this partial elegance comes at an enormously high price. It completely blurs the differences and peculiarities of different levels of the world. Moreover, Simons' 3,072 fundamental modal

[15]The following quotations are taken from a private e-mail communication.

combinations empty ontology of any explanatory power, at least for my taste. Why not say that *everything* in the world is fundamental. Simons' attempt to account for the diversity of being results in a surrender of *understanding* the ontological structure of the world. In the case of QFT, this surrender manifests itself in the fact that Simons' categorial scheme has no connection to the formalism(s) of QFT. Simons' proposal does not help to understand how the world might look according to the results of QFT. The scheme might be good for a computer-supported classification of the world—since computers do not try to understand anything—but for an ontology of QFT it seems out of place.

4.4 Conclusion

I wish to end my comment with an evaluation of the significance of the contributions of Johanna Seibt and Peter Simons for the search for an appropriate ontology for QFT. I think one of the greatest merits of both authors lies in their methodological considerations, which should be very helpful for philosophers of physics who are working on the ontology of QFT. Both authors emphasize that it is vital to have a clear notion about what the task of an ontologist is and what it is not. Seibt, e. g., stresses that it is pivotal for the ontologist to uncover hidden presuppositions in order not to be mistaken about where problems stem from and which conclusions can legitimately be drawn.

The extensive discussion of both authors about various ontological theories can be of great help in at least two opposed aspects. They limit and they widen the range of possibilities. The first, limiting aspect has an immediate impact on ongoing ontological considerations in philosophy of physics. Both Seibt and Simons lay out various arguments to the effect that the most widely-known ontological conceptions are troubled by inconsistencies. Most notably, substance ontologies are regarded by both authors as out of the game. Since substance-ontological notions usually play an important part in discussions about the particle interpretation of QFT, there is no need to say how disastrous inconsistencies of substance ontologies could be. Seibt scrutinizes this issue in section 3 of her paper. The effect of Seibt's and Simons' considerations is not only a limiting one, however. The plethora of ontological conceptions will help the philosopher of physics to find new and viable options beyond the traditional schemes.

Nevertheless, after evaluating the specific proposals for an ontology of QFT by Seibt and by Simons, as I have done in the two preceding sections, I believe that trope ontology is a more promising candidate—in particular when looking at the algebraic approach.

References

Campbell, K. (1990): *Abstract Particulars*. Oxford: Blackwell.

Finkelstein, D. R. (1996): *Quantum Reality: A Synthesis of the Ideas of Einstein and Heisenberg*. Berlin et al.: Springer.

Scheibe, E. (1991): Substances, physical systems, and quantum mechanics. In Schurz, G. and Dorn, G. J. W. (eds.): *Essays in Honour of Paul Weingartner on the Occasion of his 60th Birthday.*, Advances in Scientific Philosophy, Amsterdam, Atlanta (GA): Rodopi B. V., 215–230.

Seibt, J. (1995), Individuen als Prozesse. *Logos (N. F.)*, 2: 352– 384.

Seibt, J. (1996): Existence in time: From substance to process. In Faye, J., Scheffler, U., and M. Urchs (eds.): *Perspectives on Time*, Boston Studies in the Philosophy of Science. Dordrecht: Kluwer Academic Publishers, 143–182.

Seibt, J. (2000c): The Dynamic Constitution of Things. In Faye, J., Scheffler, U., and M. Urchs (eds.): *Things, Fact, and Events. Poznan Studies in the Philosophy of the Sciences and the Humanities*. vol. 76, Amsterdam: Rodopi., 241–279.

Simons, P. M. (1994): Particulars in particular clothing: Three trope theories of substance. *Philosophy and Phenomenological Research*, LIV(3): 553–575.

Simons, P. M. (1998a): Farewell to substance: a differentiated leave-taking. *Ratio (new series)*, XI: 235–252.

Simons, P. M. (1998b): Metaphysical systematics: A lesson from Whitehead. *Erkenntnis*, 48: 377–393.

Teller, P. (1995): *An Interpretive Introduction to Quantum Field Theory.* Princeton: Princeton University Press.

This page is intentionally left blank

Chapter 5

How Do Field Theories Refer to Entities in a Field?

Sunny Y. Auyang
Cambridge, MA

5.1 Introduction

The philosophy of physics is interesting and highly specialized. In pursuing the ever more esoteric ontologies and theories of modern physics, it runs the danger of being detached from philosophy at large. Isolation harms both sides. A large body of philosophy never benefits from the insight of modern physics or faces its criticism. On the other hand, philosophical analysis of physical theories invariably invokes many general concepts, elucidation of which could be helped by work in other areas of philosophy. Therefore interdisciplinary conferences or anthologies are most valuable.

In ontology we are concerned not only with *what there is* in the universe, we are also concerned with *what concepts we have presupposed* in talking about the things that exist. Concepts structure our understanding of the world. When physicists wonder whether the dark matter in the universe is baryonic or composed of more exotic particles, they have presupposed general concepts such as entities, substances, events, and processes. These general concepts are most important for the interpretations of physical theories because they figure in the bottom lines of arguments. They are familiar; everyone intuitively understands them. However, when we ask what exactly do the general concepts mean, and what they have presupposed, we find that they are anything but trivial. As Kant argued, we have presupposed a lot in our most ordinary discourse. Thus ontological analysis cannot proceed without an accompanying clarification of general concepts.

In this paper, I try to examine the field ontology by using some general concepts from philosophy of language, specifically, concepts involved in referring to particular entities. The approach is encouraged by the conception of gauge field theory as a language or formal framework for fundamental theories in modern physics, including general relativity and quantum field theory. This language enables us to talk about and study things far from our ordinary experience. I try to show how some philosophical results in the debate on reference help to clarify ontological issues in the interpretations of gauge field theories. Conversely, viewing the gauge field framework as a successful language for modern physics, it challenges philosophers of language to explain how their theories apply to it.

An interdisciplinary study depends on the disciplines to be bridged. The three areas considered here are ontology, gauge field theory, and philosophy of language; more specifically, individual entities, local fields, and reference. I introduce gauge field theory and philosophy of language before explaining how they illuminate each other and ontology. On the field theory side, I take as my premise that the ontology of the universe is a set of interacting fields, for example the electron and quark fields, the electromagnetic and the gluon fields, and so on. Thus I ignore strings and other claims to the basic building blocks of the universe. I also put aside questions about what kinds of fields there are and what kind of things the fields make up. Here I consider only the general nature and structure of a free matter field, say the electron field. The electron field is a very complicated whole. To bring out its fine structures, gauge field theory successfully analyzes it into smaller entities. Are these entities substances, events, bundles of energies, or something else entirely? These are the ontological questions that are philosophically important. Note that whether the field ontology consists of substances, events, or processes, we need to talk about individual substances or particular events, hence have presupposed the concepts of *individuality* and *particularity*. By clarifying how these concepts work in field theories to refer to individual entities, we gain a better understanding of entities in fundamental physical fields.

5.2 Direct and Descriptive Reference

To see how we refer to entities in gauge field theories, let us first look at how we refer generally. In our everyday speech and in scientific theories, we instinctively refer to things. We talk about Tom meeting Mary, particle 1 scattering off particle 2. However, it is not clear how we refer to particle 1 or what concepts have we used in making the reference. To clarify it is the job of philosophy of language.

Language meets reality when we refer to things in our speech. Thus reference is the most apparent bridge between linguistic and ontological studies. Consider the most common form of discourse, the subject-predicate proposition, where we pick out an entity as the subject and then describe it by a predicate. To pick out the subject of discourse is the work of *reference.* How do we do it? Suppose we pick out Daphne as our subject and describe her by the predicate beautiful. We can do it two general ways. We can refer directly and say "Daphne was beautiful", or we can pick out the subject by a description, as in "The woman chased by Apollo was beautiful". Both methods are commonly used. But philosophers of language have found that upon analysis they reveal different presuppositions and imply different ontologies. They lead to two major theories of reference: *descriptive reference* favored by Bertrand Russell (1905) and *direct reference* championed by Saul Kripke (1972). Direct reference presupposes an ontology of *individual entities* with numerical identities. Descriptive reference presupposes an ontology of *bundles of qualities* without numerical identities. If a method of reference succeeds in a tested physical theory, then we can infer the ontology the theory represents.

In direct reference, we use a singular term a to pick out the subject, often without mentioning its properties. Singular terms include proper names such as Daphne or electron 1, pronouns such she or it, and common nouns preceded by a definite article, such as the woman or that electron. Then we can ascribe properties to the entity designated by the singular term with a predicate F such as beautiful or having up-spin. Logically, a subject-predicate proposition is represented in the form Fa, "Daphne was beautiful" or "electron 1 has up-spin".

In referring directly to an entity by a singular term, we have made several presuppositions about the general nature of the entity. We have presupposed that to represent an individual, we need at least three general concepts: its *numerical identity*, its *possible properties*, and its *kind.* The notion of numerical identity implies that the individual designated by a singular term is always a unique particular or an individual, regardless of its properties. The sheep Dolly may be qualitatively identical to its clones, but the singular term designates it and only it. Numerical identity is expressed by the fact that the singular term is what Kripke called a *rigid designator.* It rigidly designates the same individual in different possible situations. Daphne was beautiful, and the name "Daphne" designates the same woman in a possible world where she became quite ugly when she grew old.

The notion of possibility is important because things usually change over time, but we refer to the same thing even as it takes on different properties. However, the range of possible situations where a rigid designator works is not totally arbitrary. When Daphne changed into a laurel tree, the name no longer designated the same woman because the woman ceased to exist. Thus the numerical identity of an entity is not absolute but relative to a kind of thing; Daphne refers to a woman, not a tree. There are many kinds. They can be divided into two general classes, kinds of thing and kinds of stuff. Stuff such as water or gold is undifferentiated. In contrast, things such apples and oranges are already individuated by their kind concepts. When we say apple, we have already presupposed some criteria of what counts as *one* apple. The criteria of differentiation constitute what Locke called *sortal concept* for apples. When we refer directly to an entity, we have tacitly assumed that the entity belongs to a kind of thing, hence we have presupposed some sortal concepts.

Together the three general concepts, numerical identity, possibility, and kind constitute the concept of *individual entities*. They are what Aristotle called "being-qua-being", this-something. In less imposing terminology, they are simply what we ordinarily call *things* and refer to directly in our everyday speech. Individual entities also occur in physical theories, for instance mass points, particles in quantum mechanics, and the local fields in quantum field theory. I will return shortly to show how the concepts of identity, possibility, and kind contribute to the representation of local fields.

Useful as it is, direct reference is philosophically controversial, for its conceptual framework comprising three general concepts is difficult to account for in predicate logic. Some philosophers opt for a simpler framework with only one general concept, where we refer to things via descriptions. The police put out a bulletin for the murderer of Jack Smith, the teacher asks about the highest peak on earth, and what is required is whoever or whatever fits the description. The idea of descriptive reference arises naturally in quantified logic, where the most general form of a proposition is $\exists x Gx$ "there is x, x falls under the predicate G", where x is the variable of quantification. For example, there was x, x was chased by Apollo. Whomever Apollo chased is our subject of discourse, and Daphne happens to fit the bill.

Most descriptions we use involve some singular terms such as Apollo or the earth, consequently they are not pure descriptive reference. The pure form of descriptive reference operates without the help of any singular term and uses only general predicates to pick out the subject matter. It occurs in, for example, Willard V. Quine's (1960) regimented language that has eliminated singular terms. Pure descriptive reference casts out a net with a predicate, and pulls in whatever the net catches as its subject of discourse. It refers to whatever that satisfies the predicate. In Quine's slogan, to be is to be the value of a variable.

As Peter Strawson (1950) pointed out, shorn of the support of singular terms, pure descriptive propositions assert only "these qualities are instantiated" but not "this entity has such qualities". Descriptive reference implies an ontology different from that of direct reference. What are referred to in descriptive reference have no numerical identity and no possibility to be different. They are not particulars or things in the ordinary sense. Rather, they are *bundles of qualities* as described by the predicates. Qualities as described are universals. Thus bundles of qualities are different from Keith Campbell's (1990) tropes or abstract particulars such as this patch of green; *this* or particularity is precisely what bundles of qualities lack.

Descriptive reference distinguishes itself from direct references in dispensing with the concepts of possibility and numerical identity. It need not consider the possibility that its subject has different properties, because a different description automatically refers to a different entity. Of importance, descriptive reference does not imply that what it refers to is unique. Apollo may have chased after many women, never mind, haul them all in as the values of the variable x. Predicates are general. By definition each predicate can have many instances, each instance can be the entity referred to. If we want to specify a single entity by descriptive reference, we must add an explicit qualification to assert that its referent is unique: $\exists xGx(\forall y(Gy \leftrightarrow y \equiv x))$, which reads "$x$ is G, and for all y, y is G if and only if y is identical to x". Of course, whether the uniqueness criterion is valid depends on the predicate G and the structure of the real world. Quine suggested that we find a unique predicate for every thing in the real world. As we will see, however, this is not always feasible.

Bundles of qualities without numerical identities do exist. Good examples are the quanta of field excitation, usually known as "particles", which are mathematically represented in the number representation of field theories. Field excitations are not particles in the ordinary classical sense of being similar to tiny pebbles, which are individual entities with numerical identities. A quantum of excitation in a field is just a chunk of energy satisfying the field's dispersion relation, $h\omega_k$, where k stands for the wavevector, spin, and other relevant quantum numbers. These values are the definite predicates for the states of the field quanta, by which we can refer to various field quanta.

A value of the dispersion relation is a predicate G, which applies generally to many field quanta. If we need to distinguish one quantum from another, we need to add an explicit uniqueness criterion. Such criteria are sometimes available, but not always. Notable examples are the quanta for fermion fields such as the electron field. Here a field quantum is an electron, and the Pauli's exclusion principle asserts that no two electrons can be in the same state. With this additional uniqueness criterion, description succeeds to refer to a single electron. The success, however, is qualified. Quanta of excitation are clearly defined only in free fields. When fields interact, as they must, field quanta become dubious entities.

Even for free fields, uniqueness criteria are not always available. This is the case for boson fields such as the electromagnetic field. Here the field quantum is a photon and a value of the dispersion relation specifies a photon's state. Many photons can share the same state. We call a laser beam coherent because it consists of zillions of photons all in the same state. Thus we can tell *how many* photons there are in a coherent state, but we cannot distinguish one photon from another. The case of photons shows the limits of the discriminative power of descriptive reference.

Descriptive reference puts the whole burden of ontology on the concept of quality. Unfortunately, even in field theories, where we are dealing with only a few kinds of thing, qualities alone are insufficient to guarantee the uniqueness of reference. The situation becomes far more desperate when we consider the infinite diversity of things that the elementary fields can combine into. In complicated circumstances, advocates of descriptive reference sneak in another notion to secure the uniqueness of entities, namely space or spacetime regions. Instead of a particular rabbit, Quine talked about the particular space region where the predicate rabbithood is instantiated. With that, he quietly changed the ontology by adding a spatial substance underneath the bundle of qualities. Instead of being merely the value of a variable, he switched to the traditional pincushion model where qualities are like pins that stick on to a bare substance or an empty spatial region.

Now we come to the interesting question. What is the ontological status of the space or spacetime regions? How do we differentiate them and refer to them? What are the predicates that can differentiate one spatial region from another? When these questions are taken seriously, we see that the conceptual simplicity of descriptive reference is deceptive. For with spatial regions, it smuggles back the concept of numerical identity that it claims to discard. This is best seen in the concrete case of field theories, where spacetime is treated explicitly.

5.3 How Does Linguistic Analysis Apply to Physical Theories?

How do linguistic considerations help to clarify ontological issues in the interpretation of gauge field theories? Consider Dirac's equation for the electron field:

$$\left(i\hbar\gamma^{\mu}\frac{\partial}{\partial^{\mu}} - mc\right)\psi(x) = 0.$$

In linguistic terminology, we can say that it encompasses a class of infinitely many sentences about the electron field. What do these sentences say? What are their subjects and predicates? What entities do they refer to? How is the reference made, directly or descriptively? What does the method of reference suggest about the nature of the entities?

We can distinguish four kinds of term in the equation: (1) the indexical variable x, usually called the spatio-temporal parameter; (2) the dynamical variable $\psi(x)$, called the local field operator; (3) the differential operation $\frac{\partial}{\partial x^{\mu}}$; (4) constants h, m, c, γ^{μ}. I will concentrate on the first two. The differential operator is paramount in cementing various local fields into an integral electron field and plays crucial roles in nonlocal effects, as discussed by Paul Teller (2000) and Holger Lyre (2001). However, its complex function and nonlocal effects are beyond the scope of this paper.

The two variables x and ψ are the chief ontological concepts for the electron field. How do we interpret them separately? In other words, how do we analyze the electron field into smaller entities? I will consider two general approaches, which I call horizontal analysis and vertical analysis.

Horizontal analysis starts with the assumptions that a single concept or variable is sufficient for reference, consequently the variables x and ψ each refers to something physical. Thus it decomposes a physical field into two layers. The layer referred to by x is most often taken to be the bare spacetime and the layer referred to by ψ an energetic coating, probably in the form of quanta of excitation. The two physical layers are related explicitly by supporting or containing, spacetime supports or contains energy. This horizontal analysis is crude. To know more about the structure of the field, we need to analyze each layer further into smaller entities and to find criteria for matching them up. The second analytic step breaks spacetime into many points or tiny regions, each of which is a bare particular. Thus we get the pincushion model of an entity, where the entity consists of two parts, a cushion that is a bare particular, on which stand various pins that are its qualities.

Bare particulars are individuals that are stripped of all features and qualities. Traditionally they are some kind of basic substance. In field theories they become empty spacetime points or regions. The notion of bare particulars has been under criticism for centuries, for its proponents can give no account of how to individuate them or refer to them. How do you differentiate one pincushion from the other, knowing that they are all totally featureless? Many arguments in the philosophy of spacetime show that bare spacetime is the quintessential stuff. It allows no criterion of differentiation because there is no way to mark the boundaries. Ordinarily, we differentiate spacetime regions by the material features in them. Devoid of matter, this differentiation fails.

Suppose we ignore the problem of differentiation and assume by brute force that there are bare particulars or empty spacetime regions. How do we refer to them? We cannot refer to them descriptively, because they are all identically featureless. If we refer to them at all, we must do so directly, which means that we have presupposed the notion of numerical identity. Actually the whole purpose of the bare spacetime regions is to perform the job of the concept of numerical identity, for it works on the intuition that no two things can be in the same space at the same time. Instead of simply acknowledge the concept that we use in thinking, it does the job of identification by the heavy ontological machinery of positing a distinctive substance. Ontology does not come free; it demands concepts for representation.

Now we can see the peculiarity of the pincushion model of entities. It breaks up an ordinary entity with identity and qualities into two entities, the cushion and the pins. Then it assigns numerical identity to the cushion and qualities to the pins. Thus it posits two distinct kinds of entity. Entities of one kind have numerical identities but no quality; entities of the other kind have qualities but no identity. We refer to them by two distinct methods, one directly and the other descriptively, and hope that somehow they match to each other. I can find no explanation of how the marriage works. We do not have enough concepts to handle the rich ontology of the pincushions.

We can get a more parsimonious ontology by analyzing the field vertically. Learning from the theory of direct reference that we need more than one general concept to refer to an entity, we interpret the operator $\psi(x)$ as the two-pronged variable for a single kind of entity, local field. Local fields are particular individuals represented by two general concepts, numerical identity represented by x and possible properties represented by ψ. We refer to the local fields directly via the variable x and describe each individual referred to by the predicate ψ. This interpretation agrees with the usual saying that x is the *independent* variable in the equation and ψ the *dependent* variable. Together $\psi(x)$ represents a single set of entities, the local fields. The electron field is a dynamical system consisting of infinitely many dynamical local fields, each identified by a value of the spatio-temporal parameter x and is a particular individual. That is it. We need not go further and decompose a local field into a cushion and the pins on it. Actually we cannot, because the decomposition will require more concepts, which we do not have.

Most if not all physical theories distinguish between dependent and independent variables, or dynamical and indexical variables. The functions and interpretations of the two variables are always different. In classical and quantum mechanics, space or position is a dynamical variable whereas time is the only indexical variable. Few people regard time as a substance, a kind of river in which matters flow. Yet time is absolute and objective because mechanics describes dynamical processes and processes are inherently temporal. A process is temporally extended and composed of many stages, which are indexed by the temporal parameter. Therefore the process has an intrinsic temporal structure. A big change in field theories is that space is now demoted from a dynamical variable to the same status as time, so that spacetime becomes the indexical variable with four components. In my interpretation of the local fields, the spacetime variable retains the interpretation of the time variable in mechanics. Spacetime is absolute and objective as an intrinsic structure of the fields. Beyond that, I refrain from attributing more substance to it.

5.4 Direct Reference to Local Fields

Earlier I said that direct reference presupposes three general concepts, numerical identity, possibility, and the sortal concept of a kind of thing. In field theory, the numerical identities of the local fields are represented by the spatio-temporal variable x. The other two concepts, possibility and kind, are combined in the notion of the state space.

Philosophers are still debating how we recognize a group of things as belonging to the same kind. Fortunately, this problem is not serious in fundamental physics, where there are not many kinds of fields. A common way in physics to delineate a kind of system is to find its *state space* that encompasses all possible states or all possible properties that systems of that kind can assume. State spaces are not physical spaces but mathematical structures or sets of mathematical entities endowed with certain structures. A state space not only contains a set of possible states but also includes the relations among the states.

A fundamental characteristic of a field is its *symmetry*, which means it is invariant under a certain group of transformations. Various quantum fields have various symmetries, mathematically characterized by various symmetry groups: e.g. the unitary group $U(1)$ of the electromagnetic interaction, or the group $SU(3)$ for the strong interaction. All the elements of a group constitute what mathematicians call its group space. The group space of a field's symmetry also serves as its state space. For instance, the state space of the electron field is the group space of $U(1)$, which looks something like a circle. Each point on the circle is a possible state of the electron field.

Gauge field theory localizes the symmetry group to each point in the field. Consequently each local field in the electron field is represented by its private state space encompassing its possible states. To see the ontological significance of the localization of symmetry, remember that the electron field as a whole spans the universe, and we want to analyze it into a set of local fields to which we can individually refer. All these local fields should belong to the same kind; they are all electron local fields, as distinct from quark local fields, with which they will interact. The symmetry group is a characteristic of the electron field as a whole. It also serves as the sortal concept or kind concept that individuates the electron local fields. The state spaces of the local fields are all mirror images of the same symmetry group $U(1)$, therefore the local fields all belong to the same kind. Because the localized state spaces are all disjoint from each other, the extent of one local state space becomes the sortal criterion of what counts as one local field. Thus the localization of symmetry explicitly demonstrates the operation of the sortal concept in individuation.

When we talk about a particular electron local field, $\psi(x_1)$, we acknowledge its kind by the structure of its state space. To distinguish it from other local fields with identical state spaces, we have its numerical identity x_1. This *conceptual analysis* should not be confused with *physical decomposition*. The state space and numerical identity do not refer to two things; they are two concepts that jointly facilitate reference to a single local field.

By itself, the field equation is only a framework for the general characteristics of the electron field. Its solutions under various boundary conditions yield definite sentences about objective states of affairs. Under a particular set of boundary conditions, the electron field is in a particular state, which means that each of its local fields realizes a particular state out of all its possible states. For instance, the local field $\psi(x_1)$ actualizes the state w_1. Its neighbor $\psi(x_2)$ actualizes another possible state w_2. Of course, the variation of the actualized states among local fields is not arbitrary but is governed by the field equation, especially the operation of the differential operator $\frac{\partial}{\partial x^\mu}$. The actualized states of all local fields constitute the actual state of the electron field as a whole. Under another boundary condition, the electron field would actualize another state. Thus there are many other possible states of the electron field. To use a philosophical terminology, there are many possible worlds for the local fields.

Significantly, through all variations of possible states, the spatio-temporal parameter x stays put. The value x_1 continues to be the identity of the same local field $\psi(x_1)$, no matter what state the local field is in or what value ψ assumes. This is because the differential operator operates on the field state ψ but not on the variable x. It only shows how the state changes as we go from one local field to another as the value of x changes. This is the mark between indexical and dynamical variables.

Linguistically, the spacetime variable x has the same function as the pronoun "it". Pronouns are variables although not the variable of logical quantification. The pronoun "it" refers to different things in different contexts, but in a fixed context, it functions like a name that designates a particular thing. In ordinary language, the context is fixed by the circumstance of discourse. In field theories, we can fix the context by choosing a particular coordinate system, which assigns a set of four numbers to each value of x as its coordinates. The coordinates x_1^μ function as the *name*

Reference	General concepts (example)	Ontology (example)
descriptive	quality ($h\omega_k$)	bundles of qualities (field quanta)
direct	identity (x); kind, possibility ($U(1)$)	concrete particulars (electron local fields)

Table 5.1 *Two theories of reference and the general concepts and ontology they presuppose.*

for the local field $\psi(x_1)$. Thus a coordinate system systematically assigns names to all the local fields. Names are tags, and we can tag the local fields because their concepts include the numerical identity as the "bulletin board" without which the tags would not stick. Furthermore, the names are rigid designators, as Kripke argued. They rigidly designate the same local fields through all possible worlds relevant to the electron field. Thus gauge field theory lays out clearly the concepts by which we directly refer to individual entities.

5.5 Summary

Ontological discussions depend heavily on the concepts that we have presupposed in talking about what there is. Table 1 summarizes the general concepts and ontologies presupposed by the direct and descriptive theories of references. The two theories are not mutually exclusive. Both particular entities and bundles of qualities occur in quantum fields. Local fields are particular four-dimensional entities to which we refer directly. The quanta of field excitation, usually known as particles, are bundles of qualities, to which we refer descriptively. Thus we need both theories of reference.

Research on reference has shown that there is no conceptual free lunch. Descriptive reference is conceptually simpler, requiring only one general concept, that of quality. Consequently it cannot handle entities and is restricted to the ontology of bundles of qualities. It is used in quantum field theories, where dispersion relations describe excitations of field quanta. However, the bundle ontology is too simplistic for field theory, for which the number representation alone is not sufficient. Furthermore, for many bundles, for instance the quanta of boson fields, it fails to individuate the bundles uniquely. To account for the uniqueness of reference, many philosophers sneak in the notion of bare particular or empty spacetime regions. The move fails to redeem descriptive reference, because the featureless spacetime regions cannot be picked out by descriptions. We need more general concepts for an ontology with particular entities to which we can individually refer.

The theory of direct reference shows that to refer to individual entities, we need the combination of three general concepts: numerical identity, possibility, and kind. All three concepts are used in gauge field theories to represent the local fields: numerical identity in the spatio-temporal parameter, kind and possibility in the symmetry group space localized to each point in the field. They show that to talk about concrete particulars to which we can individually refer, we need the complex of concepts. Therefore we should not carelessly multiply ontology and create mysterious substances simply to mirror the concepts that we use. More specifically, just because in talking about concrete particulars we have used the concept of numerical identity, we need not posit bare particulars or empty spacetime regions to answer for the concept. If we do, we will find ourselves at a loss to show how we individuate and refer to those empty spacetime regions. In short, what there is does not mirror what concepts we use to think about it.

Acknowledgements

I would like to thank Holger Lyre for helpful suggestions.

References

Campbell, K. (1990): *Abstract Particulars.* Oxford: Basil Blackwell.

Kripke, S. A. (1972): *Naming and Necessity.* Cambridge, MA: Harvard University Press.

Lyre, H. (2001): The Principles of Gauging. *Philosophy of Science*, 68: S371-S381.

Quine, W. V. (1960): *Word and Object.* Cambridge, MA: MIT Press.

Russell, B. (1905): On Denoting. In: *Logic and Knowledge*. New York: Capricorn Books, pp. 39-56.

Strawson, P. F. (1950): On Referring. *Mind* 54: 320-44.

Teller, P. (2000): The Gauge Argument. *Philosophy of Science*, 67: S466-81.

PART 2

Field Ontologies for QFT

This page is intentionally left blank

Chapter 6

A Naive View of the Quantum Field

Andrew Wayne
Concordia University, Montreal

6.1 Introduction

One of the central arguments of Paul Teller's *An Interpretive Introduction to Quantum Field Theory* concerns the extent to which QFT can be interpreted as a quantum theory about fields. An approach to QFT popular among physicists holds that QFT is a "quantum theory of fields," as the titles of two influential textbooks on the subject suggest (Wentzel 1949, Weinberg 1995). Historically, QFT began by quantizing the classical electromagnetic field (field quantization) and quantizing the state function as if it were a classical field (second quantization). Heuristic QFT is still presented as a quantized version of classical field theory. QFT contains entities, quantum mechanical operators $\Phi(\mathbf{x}, t)$, indexed by spacetime points, and in the Heisenberg and interaction pictures these operators evolve over time. The analogy with classical fields, represented by a set of scalars, vectors or tensors also varying over space and time, is a natural one. On this approach—call it the *naive view*—the value of a quantum field at a particular point is represented by the operator at that point. Teller argues that the naive view is "an entirely wrong-headed way of thinking about the subject" (1995, 93) and he proposes an alternative interpretive approach. Teller has done valuable work revealing a deep confusion in the naive view, a confusion present in much other work on the interpretation of quantum theories, about what a 'physical quantity' is in the quantum context. Nonetheless I believe that, suitably developed, the naive view *can* provide a good basis for the interpretation of QFT.

An interpretation of QFT must, inter alia, provide an account of the objects, properties and processes in the world to which the formal models of QFT might correspond. The naive view is not detailed or precise enough to provide such an account and is, at best, an interpretation sketch or framework for interpretation. For this reason, I read Teller's assertion that the naive view is "entirely wrong-headed" as the claim that any interpretation of QFT developed within the framework of the naive view—roughly, any interpretation in which field configurations are represented by field operators—will be inadequate. I think Teller's argument for that claim is based on an unduly restrictive conception of which elements of a model can be interpreted to correspond to quantities, as I'll explain shortly. However, the proof of the pudding is in the eating, and the best response to Teller's claim would be a fully-developed interpretation of QFT within the framework of the naive view. I have done some work in this direction, and I'll sketch some of my results.

6.2 What Teller's Quantum Field Is Not

Teller's argument begins with widely-held characterization of the field concept in physics: fields associate values of physical quantities with spacetime points. In a system containing a field, values of physical quantities are attributed to spacetime points, and specifying all the values of the relevant quantities at all points completely describes the field. Thus on the naive view, quantum field theory counts as a field theory only if the assignment of operators to spacetime points in a model corresponds to the values of some physical quantity. Teller argues that they do not, based on a distinction between a determinable, usually represented by a variable, and a particular value:

> A *Determinable* is a collection of properties such that anything that can have one of the properties in the collection must have exactly one of the properties. The *values* of a determinable are its individual properties. Values of a determinable are ordinarily represented by mathematical entities such as real numbers and vectors (1995, p. 95).

Teller points out that talk of fields is frequently ambiguous between field determinable and field value (or configuration). The phrase "electric field," for example, can refer to a particular configuration of a field, that is, an assignment of specific electric field values to spacetime points. But the phrase is also used to refer to the electric field more generally, the electric field determinable. Teller argues that field operators correspond most closely with classical determinables rather than classical values. Each field operator $\Phi(\mathbf{x}, t)$ has the same spectrum of eigenvalues, and this spectrum corresponds to all possible values in all possible states. Each operator encodes information about "the value of the physical quantity to which it corresponds, not for any one fixed system, but for any system that can have a value of the quantity" (1995, 98). Thus, field operators correspond most closely to field determinables, not specific field values. The field values themselves are given by expectation values for field operators in a specific state $\langle \psi \,|\, \Phi(\mathbf{x}, t) \,|\, \psi \rangle$, and these correspond to propensities for the manifestation of field-like or particle-like phenomena.

Teller develops his association of field operators with classical determinables by means of the following analogy (1995, pp. 100-101). Consider a classical field determinable F, the values of which are specific field configurations c. F can be decomposed into "component determinables" $F|_{(x,t)}$, which for a fixed point is a determinable whose values are all the possible values of $c(\mathbf{x}, t)$ at that point. It is possible to interpret the component determinables $F|_{(x,t)}$ at distinct points as distinct determinables and view a classical field as a collection of distinct determinables-at-spacetime- points. Teller dubs this the "perverse reading of the classical field" (1995, p. 99), and this is precisely Teller's reading of the quantum field. For a *fixed* $(\mathbf{x}, t)$, a quantum field operator $\Phi(\mathbf{x}, t)$ is a determinable in the same sense that $F|_{(x,t)}$ is a determinable, and for *variable* $(\mathbf{x}, t)$, a collection of $\Phi(\mathbf{x}, t)$ make up an "operator-valued quantum field" in just the way that the collection of component determinables $F|_{(x,t)}$ make up the perverse reading of the classical field.

Teller's analogy is misleading, however. Field operators indexed by distinct spacetime points are different operators with distinct physical content in a way that the various $F|_{(x,t)}$ are not. Each field operator has a spectrum of eigenvalues, and to each eigenvalue there pertains one or more eigenvectors. All field operators $\Phi(\mathbf{x}, t)$ (for all $\mathbf{x}$ and t) have the same spectrum of eigenvalues, but distinct field operators $\Phi(\mathbf{x}, t)$ and $\Phi(\mathbf{x}', t')$ have different eigenvectors associated with their eigenvalue spectrum (here and in what follows we work in the Heisenberg picture). What varies between field operators indexed by distinct points in space and time is the assignment of particular eigenvectors to the invariant eigenvalue spectrum. The spacetime variation of eigenvector phases gives us substantive physical information about the field. A different assignment of operators to spacetime points is a physically different situation. Of course, an assignment of operators to spacetime points does not, in general, fix the state of the system. The determinate state of a quantum field is given by the association of a set of quantum field operators with a specific quantum state vector. It is this relationship between operators and state vector which fully specifies physically measurable quantities in a quantum system.

6.3 The VEV Interpretation

I contend that the spacetime-indexed set of field operators *is* a field configuration—of a special sort. It is a useful fact about quantum field theory that certain vacuum expectation values (VEVs) offer an equivalent description of all information contained in the quantum field operators, their equations of motion and commutation relations. In general, a set of VEVs uniquely specifies a particular $\Phi(\mathbf{x}, t)$ (satisfying specific equations of motion and commutation relations) and vice versa. One can derive explicitly a set of vacuum expectation values from a field operator $\Phi(\mathbf{x}, t)$ and then perform the reverse task, specifying a unique $\Phi(\mathbf{x}, t)$ from a set of vacuum expectation values. Arthur Wightman (1956) first showed how operator-valued field equations can be equivalently expressed by an infinitely large collection of number-valued functions constraining relations between expectation values at different spacetime points. In general, n-point VEVs $\langle 0 \,|\, \Phi(\mathbf{x}_1, t_1) \ldots \Phi(\mathbf{x}_n, t_n) \,|\, 0 \rangle$, for all finite n, are needed to fully specify the content of a set of field operators. The central claim of my VEV interpretation of QFT is that VEVs for field operators and products of field operators in models of heuristic QFT correspond to field values in physical systems containing quantum fields.

The VEV interpretation proposes widening the concepts of field value and field configuration. An *actual* state of a physical system containing a quantum field corresponds to specific state vector/operator combination. The expectation values of any field operator or combination of field operators in a particular state can be expanded as a continuous sum over n-point vacuum expectation values. These are precisely the field values recognized by Teller. The VEV interpretation admits another type of field value, namely the complete set of VEVs itself. This field configuration can be thought of as the field state of the vacuum, although the cyclic nature of the vacuum means that this field configuration also contains substantive physical information about the field in any state. The vacuum state of the field contains vacuum fluctuations, and these can be thought of, as Teller does, as propensities for the manifestation of properties in the vacuum state (1995, 109). The vacuum state of the field also provides a seat for observable field-like and particle-like phenomena in specific (non-vacuum) states. For this reason, I suggest interpreting the complete set of VEVs as corresponding to a Lorentzian immaterial ether, rich in structure, which contributes to the production and explanation of QFT phenomena. The important point, however, is simply that the complete set of VEVs counts as a field configuration, and thus a spacetime-indexed set of field operators does as well.

We are now in a position to clarify the model-world correspondence relation at work in the VEV interpretation. In a simple model of a classical particle system, mathematical entities in the model correspond to actual physical entities in the system, and facts about the possible dynamical evolution of the model correspond to facts about the possible dynamical evolution of the system. In QFT on Teller's approach, the correspondence relation needs to be liberalized. Certain entities in the model—expectation values of field operators and products of field operators in specific states—do not correspond to actual values for physical quantities but rather correspond to propensities for the manifestation of properties when the system is in a specific quantum state (1995, p. 109). A complete specification of these propensities fixes the actual state of the system. It should be noted that a similar move is required in classical field theory when a dispositional interpretation is adopted. There too certain entities in the model—the field values described by values of field variables—do not correspond to actual entities in the system but rather correspond to dispositions for various phenomena under appropriate activating conditions. A complete specification of these dispositions fixes the actual state of the system. In the VEV interpretation of quantum field theory the correspondence relation needs to be liberalized further. Here, certain entities in the model (the complete set of VEVs) do not correspond to actual entities in the system, nor do they correspond to propensities for the manifestation of properties when the system is in a specific (non-vacuum) quantum state. Further, a complete specification of these expectation values does not fully determine the (non-vacuum) state of the system. Nonetheless, the complete set of VEVs does correspond to a configuration of the quantum field, namely its special vacuum configuration.

6.4 Conclusion

Teller's analysis nicely raises what seems to me to be a central interpretive problem: how to link elements of QFT models with bits of the physical world in some principled, constructive way. Mathematical formalism and abstract models do not wear their ontological implications on their sleeve; rather, interpretive principles play a guiding role. At the most general level, for example, an empiricist is guided by the principle that models at best correspond to certain observable aspects of the physical world. More specifically, the interpretation of a particular physical theory is guided by assumptions about what physical systems covered by that theory are like. Teller's own interpretation of QFT is guided by the determinable/value distinction quoted above. In closing, let me briefly explain why I believe this distinction is not a useful one for the interpretation of QFT.

Teller assumes, in the determinable/value distinction, that physical properties of systems must correspond directly to mathematical entities with determinate values such as real numbers and vectors. This determinateness constraint is too strict, however, even for Teller's own purposes. Teller suggests that field operators are determinables and their expectation values in specific states are field configurations. Yet the expectation values are not individual quantities; they correspond rather to a set of propensities for the manifestation of field quantities at spacetime points (1995, p. 101, fn. 4). So these expectation values don't count as field values by Teller's own definition. I suggest that Teller's contrast between determinables and particular values (or configurations) thereof is a serious oversimplification. Particular field values or field configurations may admit grades of determinateness, and "degree of determinateness" by itself is not a useful indicator of whether an element of a model corresponds to the value of a physical quantity. Field values in QFT range from completely determinate quantities (e.g. when the state of the system is an eigenstate for the corresponding observable) to quite indeterminate ones (e.g. the configuration of the field in the vacuum, understood as an immaterial ether).

References

Teller, P. (1995): *An Interpretive Introduction to Quantum Field Theory.* Princeton: Princeton University Press.

Weinberg, S. (1995): *The Quantum Theory of Fields.* Cambridge: Cambridge University Press.

Wentzel, G. (1949): *Quantum Theory of Fields.* Translated by C. Houtermans and J. M. Jauch. New York: Interscience.

Wightman, A. S. (1956): Quantum Field Theory in Terms of Vacuum Expectation Values, Physical Review 101, 860-66.

This page is intentionally left blank

Chapter 7

Comments on Paul Teller's Book, "An Interpretive Introduction to Quantum Field Theory"

Gordon Fleming
Pennsylvania State University

7.1 Introduction

I will begin by talking about other books. Among the current books on QFT written by physicists there are some interesting contrasts. First consider the book "Local Quantum Physics" (1992) by Rudolf Haag, a leading mathematical physicist. This book is characterised by very careful definitions of the mathematical concepts employed and (sketches, at least, of) rigorous proofs of the theorems enunciated. Many of the definitions and theorems declared in this book are essential for carrying out the calculations of transition amplitudes for realistic quantum field theories (QFT) that are more or less directly comparable with experiment. But such calculations are absent from it. Some of the reasons for the absence is that, in some instances, additional theorems, as yet unproved, would need to be established to mathematically justify these calculations or, in other instances, the calculations, 'though rigorously doable, are dauntingly complex when so executed. Of course the book in question is aimed at the precise formulation of QFT and its structure. Applications are not the primary concern. We hear a great deal about unitarily inequivalent representations, the questionable status of Fock space and particles, Haag's theorem, the strange non-local behaviour of the vacuum state captured by the Reeh-Schlieder theorem and the local algebraic approach to QFT. We also learn that to the date of its publication *no* non-trivial examples of interacting QFTs in 4-dimensional Minkowski space-time had been rigorously demonstrated to be internally consistent! Not one.[1]

[1] I have recently been informed that *one* such model has now been proven consistent. I have not yet tracked down the technical details to assess the significance of the example.

Next consider the 2-volume magnum opus, "Quantum Field Theory" (1996), by Steven Weinberg, a leading elementary particle theorist and a principle architect of the unification of quantum electrodynamics and weak interaction field theory. This book is not mathematically rigorous. It combines mathematical deduction with heuristics and physical intuition. It works throughout with Fock space and never mentions unitarily inequivalent representations or Haag's theorem or the Reeh-Schlieder theorem. Most importantly, it is focused on the development of efficient, powerful modes of calculation of the transition amplitudes of realistic QFTs that can be compared with experiment. So far those comparisons lead to the most precise corroborations of theory to be had anywhere in physical science.

Now the overwhelming majority of elementary particle theorists and experimentalists will read and study this last book and others like it. And most of them will never read the first book or others like it. Those books are studied by a comparatively small army of aspiring mathematical physicists who want to push for the rigorous securing of QFT rather than the expansion of its domain of successful application. My reason for spelling out this contrast is that the mode of presentation of QFT in these two books, and the experience of the student in plowing through them, and the battery of learned material once the study is completed, is so different that one may argue that it is a tenuous claim that both books are about the same QFT.[2] It is then important to point out that notwithstanding comment by Paul Teller on Haag's theorem and unitarily inequivalent representations etc., Tellers book is about the version of QFT that the working elementary particle physicist learns and uses and not about the QFT that the smaller community of mathematical physicists struggle with. Since most of the published reviews of Tellers book have, in their critical function, focused more on the aspects of QFT that exercise the mathematical physics community, I have decided to concentrate my remarks on the aspects of QFT that, it seems to me, Teller chose to emphasize. The working stiffs QFT, as it were. And most of my comments will be concerned with the highly interpretive 2^{nd} and 5^{th} chapters of Paul's book.

[2]Probably the most significant book to present the case, by example, for the equivalence of the two forms of quantum field theory is the monograph on constructive quantum theory by Glimm and Jaffe (1987).

7.2 Teller on Quanta

In chapter 2 we find a valuable introduction to and analysis of some of the peculiar features of quantum particles and many particle states. But I will argue that Teller goes too far in some of the contrasts he draws. I will challenge the wisdom of his suggestion to abandon the term "particle" in favor of "quanta". I will argue that his vociferous campaign to oust the "excess formal structure" of the labelled tensor product Hilbert space formalism (LTPHSF) on behalf of the ascendency of Fock space is itself excessive. And finally I will claim that he places too much weight on the analogies he seeks to help us understand the fact that all indistinguishable quantum particles are either bosons or fermions.

Teller is at pains to emphasize the profound differences between the quantum and classical versions of the particle concept, and rightly so. But when his concern for the misleading effects of our classical associations with the term "particle" leads to the suggestion of abandoning the term and replacing it with "quanta" I fear he risks elevating the communication barrier between physicist and philosopher. The currency of "particle" among quantum physicists is extremely broad and "quanta" is used more widely to denote the minimal units of change or aggregation of whatever categories are permitted by the theory. Quantum particles are just a particularly important kind of quanta.[3]

It is true, however, that quantum particles are quite different from classical particles and Teller's denial of primitive thisness to the former is a valuable characterisation of some of that difference. On the basis of the absence of primitive thisness he then argues for the undesirability of the excess formal structure of the LTPHSF for expressing many particle states. He suggests that within that excess formal structure a vestige of primitive thisness can manifest itself. If all the particles present were distinguishable then every vector in the LTPHSF would correspond to a physically possible situation. When some particles are indistinguishable from one another then only appropriately symmetrized or anti-symmetrized subspaces of the LTPHSF contain physical state vectors. But LTPHSF is the natural

[3]Since writing this a few years ago I have become more sensitized to the widespread problems of understanding that are traceable to the classical associations that cling to the term "particle". Retaining my stated aversion to the replacement by the already overused term "quanta", I find myself toying with the suitability of the replacement term "quanton". Any takers?

formalism that emerges from a naïve move from single particle to many particle QM. What to do? Teller is rather strident in the claim that something needs to be done. We can't have this excess formal structure lying around devoid of physical interpretation! As good philosophers we are duty bound to ask why it is there. Fortunately the Fock space approach, very nicely presented in chapter 3, will save the day by erecting a structure in which the offending uninterpreted state vectors never appear.

I can't quite shake the feeling that Teller gets away with this stridency only by virtue of the ready availability of the Fock space solution. Surely excess formal structure, per se, is not inherently so objectionable. For example, it seems that gauge fields in both classical electrodynamics and the Standard Model in QFT qualify as excess formal structure. They are not directly susceptible of physical interpretation. Only particular combinations of their derivatives and/or integrals are. And even if the field theories in question could be reformulated without the gauge fields (electrodynamics clearly can; Mandelstam (1962)) at least some of the resulting structure would be rather cumbersome and obscured by the reformulation. In particular it is only through the use of gauge fields that we know how to formulate these theories as *local* field theories. So excess formal structure does not seem to be such a bad thing in *this* context.

Teller also uses the occurrence of states with indefinite numbers of quantum particles to support Fock space over the LTPHSF. States with indefinite numbers of quantum particles, i.e. superpositions of states with *distinct* definite numbers of quantum particles, are a *sine qua non* of QFT. But for this purpose the appropriate comparison is not between Fock space and a single tensor product Hilbert space but between the former and the *direct sum* of all possible tensor product Hilbert spaces. In this sum indefinite particle number states are as common as in Fock space. To be sure the excess formal structure is carried along and, in a sense, multiplied, but it is just more of the same as before.

Towards the end of chapter 2 Teller tries to provide some account of how we are to think about those indistinguishable particles that appear only in fully symmetrized states, the bosons, and those that appear only in fully anti-symmetrized states, the fermions. He provides the analogy of deposits and withdrawals in a checking account or of travelling pulses along a taut rope for bosons but is at a loss to find a similarly effective analogy for fermions. He seems to suggest that this absence of a good analogy is a defect in our understanding to be remedied by future thinking on the matter. But whether a good analogy from the domain of common experience can be found for anti-symmetrized states or not there is a caution to be sounded here. Such analogies can not deepen our *understanding* of fermions or bosons and may mislead if we take them more seriously than as pedagogical devices to help us over the barrier of unfamiliarity.

Related to this issue is Teller's remark, earlier in the chapter, that all the states in the LTPHSF may be deserving of a physical interpretation because we know what the world would be like if all the states were physically possible. But he doesn't tell us *what* it would be like. I will tell you. It would be as if every collection of indistinguishable quantum particles were sometimes bosons and sometimes fermions and, in the case of collections with more than two indistinguishable particles, sometimes particles of more exotic symmetry types.[4] The probabilities for the various options would vary from state to state but the quantum particles would never behave as if they were *neither* bosons *nor* fermions *nor* particles of the more exotic symmetry types. This result requires only that all the operators representing observables are symmetric functions of the dynamical variables referring to individual quantum particles and that states are represented by state vectors. But the former is what we *mean* by saying the quantum particles are indistinguishable. Now if we didn't add the restriction of electrons, say, to purely anti-symmetrized states but allowed indistinguishable electrons to behave like other symmetry types as well as fermions then the classical macroscopic world would be profoundly different. The Pauli exclusion principle, the historically first version of the fermionic restriction, would not hold. Consequently there would be no stable atoms and molecules. No periodic table of the elements. No crystal structures. Very probably no macroscopic solid individuatable objects. Thus no concept of primitive thisness would have evolved in our brains. Indeed our brains would probably not have evolved. I think it may be asking too much to hope to find a concept, among those we use for dealing with the macroscopic classical world, which will provide a good analogy for a counterintuitive feature of the microworld that is, itself, crucially responsible for broad qualitative aspects of the macroworld. We welcome analogies, when available, that will help us become familiar with the counterintuitive microworld. But

[4]The exotic symmetry types give rise to what has been called "parastatistics" in the literature (Messiah et al, 1964). An early and still very accessible introduction is provided by Dirac in his (1958). See sections 55 and 56.

understanding can only be expected to run in the other direction, i.e. understanding the less fundamental in terms of the more fundamental. In the present context QFT is the fundamental level and I felt at times that Teller wanted macroworld analogies for more than familiarizing purposes.

7.3 Teller on Fields

Now I turn to Ch. 5. It is here that I find my deepest disagreements with Teller and it is here that he is spelling out the arguments for what on p.9 he called the single most important point he wanted to make in the book. Namely that one should not think of QFT as being *primarily* a field theory. Thus the name "Quantum Field Theory" is a serious misnomer. I think this position is wrong and that Teller fails to make his case for it. I will argue that Teller's case rests upon (1) an excessively conservative attitude towards the uses to which the term "value" can legitimately be put, (2) a technical lapse in characterising the field operators associated with space-time points and (3) a quantitatively inadequate conception of how one can extract the field theoretic content that Teller agrees exists in QFT.

On p. 95 Teller defines the concept of a *determinable* and its *values*: *"A determinable is a collection of properties such that anything that can have one of the properties in the collection must have exactly one of the properties.*

The values of a determinable are its individual properties. Values of a determinable are ordinarily represented by mathematical entities such as numbers and vectors."

The first two sentences of the definition allow a broad construction since what can constitute a property is not spelled out. The third sentence reads like an aside telling us how values and therefore properties in this context are "ordinarily" represented mathematically. It is not clear how binding this third sentence is intended to be. There shortly follows a definition of a *field configuration for a collection of determinables* as a specific assignment of values in the collection that constitutes the determinable to each space-time point. On the next page we read that determinables and their values are "inherently classical notions" used to characterise classical fields and that a major task is to find reasonable analogues to them in quantum theories. *This* suggests that the elusive third sentence is indeed intended to be binding. But even so the *inherently classical* nature of determinables and their values escapes me. For even if we insist that these values be represented by sets of numbers, all that is required to represent Hilbert space operators is access to infinite sets of complex numbers, i.e. the matrix representation of the operator in a chosen basis.

Teller argues, however, that the (self adjoint) operators that are assigned to space-time points in QFT can, at best, be analogized to entire determinables rather than the values of determinables. The reason is that the operators determine an eigenvalue spectrum (a set of possible number values) and under unitary evolution from space-time point to space-time point the eigenvalue spectrum doesn't change. But such operators also determine the *eigenvectors* that are associated with the eigenvalues and these eigenvectors *do* change from space-time point to space-time point. So if the determinable is taken as the set of all possible eigenvector associations with a fixed set of eigenvalues then the operator assigned to a space-time point is equivalent to assigning a value of the determinable to that space-time point. To that extent the analogy with a classical field holds.

It is still the case, however, that unlike a classical field configuration these values of determinables consist of numbers and vectors that only define the operator. They do not reflect the physical state of the system. For that one must bring the state vector to bear. When we do that Teller feels more comfortable that we are approaching the true quantum analogue of the classical field. And on p.101 he seems to claim it is the expectation value of the field operator in the quantum state of the system that provides the appropriate analogue. My problem with this identification is that these expectation values severely underdetermine the quantum state and its relation to the operator field. To avoid, along these lines, this underdetermination one must consider the expectation values of *all possible products* of the operator fields with each factor evaluated at arbitrarily chosen space-time points. Then, finally, one has a complete set of number valued field-like structures from which all physical quantities can, in principle, be calculated! But now these are no longer fields associated with *single points* of space-time but, rather, with arbitrary *sets of points* of space-time. Perhaps Teller would construe this feature as yet another reason for denying genuine field-like character to the theory. For myself, however, this simply points out the infinitely richer structure of QFT over CFT and the highly non-local character of quantum correlations. These require n-point *number valued fields* for all n and can be compactified to a formalism with just 1-point fields only if one admits *operator valued fields.*

Throughout these pages Teller repeatedly makes temporary moves into a mode of discourse in which the structure and predictions of the theory are described in terms of evolving potentialities and propensities for various manifestations to occur. That happens to be a mode I favor and when Teller is in that mode I find myself purring. It is only when he comes out of that mode and feels the need to castigate the retention of classical terms like "particle" and "field" because of their misleading influences that he and I part ways. For me it seems quite sufficient to encode the differences between the quantum world and the classical world (and those differences *are* profound) by simply retaining the prefix "quantum". Quantum particles are not classical particles but they are closer to classical particles than to classical fields by virtue of being localizable and, if stable, satisfying an energy-momentum relationship parameterized by a definite rest mass. Quantum fields are not classical fields but they are closer to classical fields than to classical particles by virtue of being distributed over all of space-time and not having a rest mass. Quantum fields and quantum particles are more closely related than classical fields and classical particles. A physical state for a set of interacting quantum fields can be described, at least temporally asymptotically and perturbatively, as a system of types of quantum particles of temporally variable and indefinite number and other properties. A sensible reason for calling such a system a quantum *field* system rather than a quantum *particle* system is that the quantum particles are so much more ephemeral as Teller notes at the end of this chapter. *They* come and go like the wind and may not be present in *definite* numbers at all. The quantum fields, on the other hand, are identified at the outset and remain fixed in type and number throughout. It is *they* that persist!

The last half of this chapter is taken up with a brief assessment of the relationship between "Fields, Quanta and Superposition" followed by a discussion of specifically field-like states and phenomena that occur in QFT. I had no major problems with these sections modulo their dependence on the first half of the chapter which I have just criticized. But I found the first of the two sections a bit obscure and both suffered from terminological oddities.

In the first section, on p.104, Teller distinguishes "fields in the weak sense", for which field configurations are *not* linearly superimposable, from "fields in the strong sense" for which superposition *is* possible. But this amounts to identifying free, non-interacting fields as "fields in the strong sense" since it is only free fields for which superposition of their configurations leads exactly to new possible configurations. Approximate superposition in the presence of weak interactions is common but that qualification was not mentioned. Exact superposition of quantum *states* is always possible regardless of interactions since those states are represented by vectors in a linear vector space. But not so for field configurations and to call free fields "fields in the strong sense" seemed odd.

In the last section Teller discusses briefly the interesting cases of coherent states (the eigenstates of quantum particle annihilation operators), vacuum fluctuations (to show that the absence of quantum particles does not mean that nothing is happening) and Rindler-quanta—the conjectured quanta (not quantum particles) that some theoretical models indicate an ordinary quantum particle detector would respond to if it were accelerating in the vacuum. In considering these Teller emphasizes the relationship between vacuum fluctuations and Rindler-quanta but ignores the relationship between coherent states and Rindler-quanta. This leads him to conclude on p.111 that "—in [the vacuum state] no Rindler-quanta actually occur, so the status of [the vacuum state] as a state completely devoid of quanta is not impugned." But this goes too far. The vacuum state is not an eigenvector of the Rindler-quanta number operator and that justifies saying no Rindler-quanta *actually* occur (in keeping with the linguistic conventions of the book to this point). But the vacuum expectation value of the Rindler-quanta number operator is not zero and grows with the acceleration of the detector. So it is not in keeping with the linguistic conventions of the book to say the vacuum is *completely devoid* of quanta. In fact when the vacuum state vector is expressed as a linear superposition of Rindler-quanta-number eigenvectors it looks very much like a coherent state expressed as a linear superposition of ordinary quantum-particle-number eigenvectors. And Teller would not say that a coherent state was "completely devoid" of ordinary quantum particles. I think this point has been made in an earlier review.

7.4 Conclusion

Pretty uniformly I found myself happiest in those comparatively technical chapters where Teller is presenting the structure of the theory. He does that very well! Appropriately enough, I was most provoked and questioning when he was interpreting that structure. As a professional physicist I bring a lot of preformed conceptions to this reading and am perhaps nearly out of my depth in the turbulent waters of philosophical interpretation. But provocation of the reader to the articulation of divergent views which might otherwise lie unformulated is a worthy effect of a philosophical work. Without reservation I hold this book to be a valuable introduction to the interpretive problems of the working physicists QFT.[5]

References

Dirac, P. A. M. (1958): *The Principles of Quantum Mechanics* (4th edition). Oxford: Oxford University Press.

Glimm, J. and Jaffe, A. (1987): *Quantum Physics.* Berlin: Springer.

Haag, R. (1992): *Local Quantum Physics.* Berlin: Springer.

Mandelstam, S. (1962): Quantum electrodynamics Without potentials. *Annals of Physics*, 19:1-24.

Messiah, A. M. L. and Greenberg, O. W. (1964): Symmetrization postulate and its experimental foundation. *Physical Review*, 136B:248–267.

Weinberg, S. (1996): *Quantum Field Theory* (Vols. 1, 2). Cambridge: Cambridge University Press.

[5]Rereading this article, which was composed shortly after the appearance of Teller's book, I am most struck by one failure on my part. Nowhere do I comment on the change in attitude that had arisen in the physics community concerning the status of non-renormalizable interactions and the doubtful fundamentality of local quantum field theory. As a partial remedy for this lapse I refer the reader to Steven Weinberg's (1996) magnum opus and the index heading, 'effective field theories'.

Chapter 8

So What *Is* the Quantum Field?

Paul Teller
University of California at Davis

8.1 Introduction

Andrew Wayne and Gordon Fleming have offered trenchant critical comment on my (1995),[1] chapter 5 treatment of the so-called "operator valued quantum field" (OVQF) where I criticized the view that when quantum field theory presents us with a space-time indexed set of operators we can think of these as representing the field values of a determinate field configuration over space-time.[2]

On an important point I completely agree with both of them. On page 101 one reads that "The natural and smoothly working candidates [for field configurations of the quantum field] are the expectation values $\langle\phi|\Psi(\mathbf{x},t)|\phi\rangle$. If only I had written "and expectation values of all possible products of the space-time indexed field operators with each factor evaluated at arbitrarily chosen space-time points", which are needed for a complete characterization of a specific state. But this acknowledgment does not begin to satisfy Wayne and Fleming. I wanted to think of the space-time indexed field operators, taken collectively, as more like a classical determinable than like a classical field configuration. They insist that so doing "...is based on an unduly restrictive conception of which elements of a model can be interpreted to correspond to quantities...." (Wayne 2002) and that the case "...rests upon...an excessively conservative attitude towards the uses to which the term 'value' can legitimately be put." (Fleming 2002) I had written that

[1] Unless otherwise specified, cited page numbers are to this work.

[2] Fleming also offers many other comments on my (1995). For the most part I accept his other comments, remaining differences probably to be ascribed to differences in taste and emphasis, such as the appropriateness of terminology. It is particularly useful to have the reactions to the work by a philosopher of a philosophically well informed and acute physicist.

"Values of a determinable are ordinarily represented by mathematical entities such as real numbers and vectors" (p. 95). But, object Wayne and Fleming,, operators can be represented in matrix notation. So why should not operators, indexed by the space-time parameter, count as a specific field configuration? The (1995) presentation was misleading, Wayne and Fleming suggest, by thinking of operators only in terms of their associated eigenvalue spectrum, which is the same for all the space-time indexed operators of the OVQF. But, of course, these operators differ, specifically they differ by associating varying eigenvectors with a given eigenvalue. The configuration of this association could, in principle, be something quite different from what in fact obtains. Consequently the OVQF is really looking like a specific field configuration, associating a variation in physical properties, albeit very abstract and complex ones, with the variation in space-time location. Indeed, the configuration given by the OVQF carries a lot of information about the physical world, describing facts which could have been otherwise. So thinking of the OVQF as a determinable rather than a determinate field configuration looks to be misguided.

Wayne also puts the point in terms of vacuum expectation values (VEVs), that is, vacuum expectation values for the product of field operators at two or more distinct points. On the one hand, these collectively carry exactly the information of the OVQF. But on the other hand the VEVs are precisely the sort of generalized field values—characterizing relations between space-time points and not just properties at points taken one at a time - which I am acknowledging to be needed for any adequate way of thinking about the quantum field. So again, the status of the OVQF as a field configuration (in the foregoing generalized sense) appears to be vindicated.

I'm going to explain how I can agree with the gist of all of this—except the conclusion.

Wayne and Fleming's comments are cogent in application to those arguments of (1995) that they explicitly discuss. But they both neglect to comment on one of the arguments given there which, however, needs qualification and in any case clearly failed to communicate what I was grasping for. So I am happy to have this opportunity to clarify both the position and the arguments. Doing so will require some distinctions and appeal to some admittedly imprecise notions. However, if the reader will bear with these, I am hopeful that by the end we will at least have some better ways of understanding what is at issue in these discussions.

8.2 Giving a Physical Interpretation to the OVQF

To help set ideas let's keep in mind a familiar contrast case, the classical electromagnetic field. One might suggest that this case differs from that of the OVQF in that, in the case of the electromagnetic field, a field configuration is comprised by an assignment of a manifest value to each point while the OVQF configuration involves only an assignment of a complex disposition to each point. But, of course, this can't be the contrast: A completely determinate value of the electromagnetic field at a point is still a complex disposition, in particular a disposition to display various motions of a test particle placed at the point. The dispositions involved in the OVQF configuration may be at a higher level of generality or abstraction than those of the electromagnetic field, but at most we would have a difference in degree, not a difference in kind, among these cases.

Instead I want to suggest a striking difference: In the case of the electromagnetic field a determinate field configuration is physically contingent in the sense that collective alternative values are physically possible, but the OVQF configuration is physically necessary—there are no physically possible alternatives. This bald statement will have to be qualified, and ultimately it may well not stick in all cases. But I want to begin with this statement in order to set ideas and in the hope that, after the qualifications, we will still be able to see an interesting and important difference in the cases we are considering.

8.2.1 *Some Illustrations*

To make out the way in which I want to use the contrast between physical possibility and necessity, let me illustrate how these contrast in a simple case not involving a field configuration. Consider Newton's gravitational force law for the mutual attraction experienced by two massive particles: $f = g\frac{m_1 m_2}{r^2}$. If we read 'm_1' and 'm_2' as referring to the determinate values of the masses of particles 1 and 2 and 'r' as referring to the value of their actual spatial separation, the formula specifies the attractive force between the two particles, a fact which is physically contingent in as much as 1 and 2 could have had different masses and/or a different spatial separation, in which case their mutually attractive force would have been correspondingly different. But in other respects the information specified by $f = g\frac{m_1 m_2}{r^2}$ is not contingent: In a Newtonian world it is physically necessary that the exponent on r be 2; and the value of g is also physically necessary, or at least it is not physically contingent in the way that the masses of and separation between 1 and 2 are contingent. True, the kind of physical necessity involved, in a Newtonian world, for the value of the exponent and the value of g might be importantly different. But for the moment I want to suppress that kind of contrast and set out ideas with a, perhaps crude, two sided contrast between physical contingency and necessity.

We can spell all of this out one step further by rewriting the gravitational law as $F = g\frac{M_1M_2}{R^2}$. Here read 'M_1', 'M_2', and 'R', not as referring to determinate values for the masses and separation of the two particles, but as parameters to be filled in by determinate values and so representing the quantities, the mass of the first particle, the mass of the second particle, and their separation. So read, $F = g\frac{M_1M_2}{R^2}$ still gives a lot of information about the physical facts in the simple Newtonian world we are considering. It tells us what the form of the interaction will be, and it covers all the contingently possible cases which arise when determinate values are written in for M_1, M_2, and R. But the facts covered by $F = g\frac{M_1M_2}{R^2}$ itself are physically necessary, furthermore subsuming all the contingent possibilities which arise by specifying the values for M_1, M_2, and R.

Now let's consider some completely determinate field configuration of the classical electromagnetic field. If this configuration in fact obtains, this will be a physically contingent fact: We say that, while one field configuration obtains, alternative configurations were physically possible. So the determinate configuration of the electromagnetic field is analogous to the facts given by $f = g\frac{m_1m_2}{r^2}$ for some determinate values, m_1, m_2, and r. The force specified by $f = g\frac{m_1m_2}{r^2}$ could have been different if m_1, m_2, and/or r had been different, though if m_1, m_2, and/or r had been different, the alternative force would, of physical necessity, still have had to satisfy the form of $F = g\frac{M_1M_2}{R^2}$, which specifies the physical necessary aspects of the situation. In the analogy, there are many contingent alternatives to a determinate configuration of the electromagnetic field, while the physically necessary facts, in common to all physically possible configurations, are given by Maxwell's equations.

8.2.2 *How to Think About the OVQF*

Turning now to the OVQF, I want to resist thinking of it as a determinate field configuration which is one among many contingently possible alternatives in the way that a determinate configuration of the classical electromagnetic field is one among many contingently possible alternatives. Instead the OVQF carries information about the world which is physically necessary in the kind of way that such information is carried by $F = g\frac{M_1 M_2}{R^2}$ and by Maxwell's equations. The OVQF is, if you like, *formally* a field configuration. And as Wayne and Fleming emphasize, it carries information about the world. But this kind of information is not like the physically contingent information carried by one, as opposed to another, configuration of the classical electromagnetic field. Instead it is more like the physically necessary information carried by Maxwell's equations.

Why should we think of the OVQF in this way? An argument I gave in (1995, pp. 101-3), and which I will qualify below, goes like this. Let $g(x, t; k)$ be a complete set of solutions to a classical linear field equation. Then the form for any solution (sometimes referred to as the "general solution") is $\Psi(\mathbf{x}, t) = \int d^3q(\mathbf{k})g(\mathbf{x}, t; \mathbf{k})$ where the $q(\mathbf{k})$ can be set arbitrarily, up to some very general continuity requirements. This general form for any determinate solution carries exactly the information of the field equations. The field equations hold if and only if any physically possible world is one with one or another of the field configurations conforming to this general form. In this way the general form specifies what is physical necessary by subsuming the full range of contingency in what is physically possible for determinate field configurations.

Field or second quantization drastically reorganizes the way physical necessity and contingency are encoded in the field representation . Quantization reinterprets $q(\mathbf{k})$ as an operator valued function, $\hat{q}(\mathbf{k})$, satisfying canonical commutation relations, so that $\hat{\Psi}(\mathbf{x}, t) = \int d^3\hat{q}(\mathbf{k})g(\mathbf{x}, t; \mathbf{k})$ is a quantum mechanical dynamical variable. To physically interpret $\hat{\Psi}(\mathbf{x}, t)$ one must specify how the value for each relevant physical quantity will be represented by eigenvectors and eigenvalues. But once this conventional slack has been taken up, $\hat{\Psi}(\mathbf{x}, t)$ represents, not the values and expectation values of physical quantities, but the way in which such values can vary over space and time. Which of the many contingent configurations of expectation values actually occurs is represented only by combining $\hat{\Psi}(\mathbf{x}, t)$ with the state vector that specifies the physical state that actually occurs. In this way $\hat{\Psi}(\mathbf{x}, t)$ specifies what is physically necessary by covering the full range of physical possibilities delimited by the field equations after interpretation in terms of a quantum mechanical dynamical variable as specified by the canonical commutation relations and the overall choice of representation of physical values by eigenvectors and eigenvalues.

Wayne and Fleming had objected to my (1995) statement of the foregoing argument first by commenting that one can think of $\hat{\Psi}(\mathbf{x},t)$ as a space-time dependent association of eigenvectors with eigenvalues. They then argued that there are many different possibilities for this association so that a choice of one of them counts as some determinate $\hat{\Psi}(\mathbf{x},t)$ that consequently counts as one among many possible field configurations. This is, of course, correct, as long as 'possible' is here read as 'logically possible'. But, as explained in the last paragraph, once all the convention-fixing choices have been made, only one $\hat{\Psi}(\mathbf{x},t)$ is *physically* possible, and this $\hat{\Psi}(\mathbf{x},t)$ represents what is common to all the various more detailed physically possible configurations of expectation values, so that $\hat{\Psi}(\mathbf{x},t)$, common to all these physical possibilities and characterizing the pattern of variation in space-time of each and every one of them, itself counts as physically necessary. I propose to summarize the foregoing in the slogan, "$\hat{\Psi}(\mathbf{x},t)$ is the unique quantized solution to the field equations as constrained by the imposed canonical commutation relations."

Before getting to the qualifications I want to present a very general way of organizing the foregoing considerations. P. 95 specified that

> *A determinable is a collection of properties such that anything that can have one of the properties in the collection must have exactly one of the properties. The values of a determinable are its individual properties.*

But what is the force of the 'can' and the 'must' in this characterization? In the case of the determinable, mass, it is plausible that being a physical object includes the condition that each physical object has (at a fixed time) some completely determinate mass. This suggests reading the modal locutions as describing logical possibility and necessity.[3] When so read I

[3]Should the relevant modality be described as logical or metaphysical? On the one hand one might say that it is part of what the expression 'physical object' means that something to which it applies must have exactly one mass, which would support the characterization of the relevant modality as logical. On the other hand it might be better to speak here of metaphysical possibility and necessity if one is thinking that it is part of the essential character of a physical object, part of what it is to be a physical object, that it have exactly one mass. As nothing which follows should depend on how these details are worked out I will here use 'logical' to characterize the relevant modality.

will speak of *L-determinables.* Mass and shape are plausible examples of L-determinables.

However we could take the modalities involved in the characterization of determinables to be those of physical possibility and physical necessity, that is the possibilities and necessities allowed and imposed by the laws of the theory. On this reading I will speak of *P-determinables*: Anything which as a matter of physical possibility could have one of the values of a P-determinable, as a matter of physical necessity has exactly one.[4] It is easy to show that anything which is an L-determinable is a P-determinable, but the converse need not hold.

To apply the idea of a determinable to fields let's think of individual determinate field configurations as the values of a field determinable, where the field determinable is the full collection of possible alternative determinate configurations. When Wayne and Fleming insist that the OVQF is just one of many logically possible assignments of operators to the space time points, and so should count as a value of a field determinable, I can agree as long as we agree that it is an L-determinable that is in question. But it's an odd case because only one value of this field determinable is physically possible. The value which actually occurs does so of physical necessity. Moreover, this one occurring value subsumes all the physically contingent possibilities for physical states: By combining the OVQF with one or another contingently occurring physical state we get one or another complex arrangement of expectation values for occurrences at various places and times and also the correlations among these. In this respect the OVQF, itself the uniquely occurring value of the L-field determinable, is also a P- (but not an L-) field determinable that covers the collection of alternative physically possible configurations of expectation values.

The argument transfers also to Wayne's presentation of the material in terms of VEVs. The vacuum expectation values are the expectation values of what one would find if one were to measure various quantities in the vacuum. But many of them, or their linear combinations, are also interpretable as expectation values for various quantities in non-vacuum states, since non-vacuum states can be expressed in terms of applying field operators to the vacuum state. Thus the VEVs collectively provide a com-

[4]p. 95, footnote 1 opts for 'physical' as the reading of the relevant modalities. From the point of view of the considerations which follow, that was, at the very best, a lucky guess.

pendious list of all the physical possibilities allowed by the field equations, such that, of physical necessity, exactly one of these possibilities obtains.

8.2.3 *The Qualifications*

I summarized the account with the slogan that $\hat{\Psi}(\mathbf{x},t) = \int d^3\hat{q}(\mathbf{k})g(\mathbf{x},t;\mathbf{k})$ is the unique quantized solution to the field equations satisfying the imposed commutation relations. But, formally, any unitary transformation of this solution will again be a solution satisfying the commutation relations. In addition there will be unitarily inequivalent solutions, even unitarily inequivalent Fock space solutions. The more accurate statement covered by the slogan appealed to the qualification, "once all the convention fixing choices have been made." One may sensibly question whether unitarily equivalent, and especially unitarily inequivalent, alternatives correspond only to alternatives in conventions.

What counts here is not the accuracy of the claim of uniqueness of solution to the field equations but the claim that the OVQF holds of physical necessity. It was in aid of the latter claim that the former claim was made. If there now appear to be exceptions to the uniqueness claim we need to look at whether, or in what way, these exceptions go beyond convention fixing choices in ways that confer some kind of physical contingency on the OVQF. The argument is also complicated by the circumstance that the stark contrast between logical and physical modalities is, most plausibly, also an oversimplification. Recall the contrast between the way g and the exponent of R are physically necessary in $F = g\frac{M_1 M_2}{R^2}$. Rather than suppose that there is an objective, absolute, and context independent distinction between logical and physical modalities, I suspect that the contrast is set in a way that depends on theory. For a consideration that is much more general than the example from Newton's law of gravitation, consider that what counts as the logical modalities can be set by how we set up our state space, with the physical modalities then being set relative to the state space by specification of the laws which determine the allowable trajectories in the state space. There will often be in principle, and probably at least some times in practice, some trade-off between how we specify the state space and the allowable trajectories within the state space.

The upshot of these complications is that we should not expect some completely general way, independent of the details of the theoretical context, for giving a physical interpretation of the OVQF. Instead we have an, admittedly oversimplified, way of thinking physically about the OVQF and its similarities and differences from classical field configurations. We then look at individual cases and refine that simple blueprint, adjusting it to the theoretical context in question. But, I suggest, looking at a few examples will support the conclusion that the oversimplified account generally does exceedingly well.

Let's begin with the worry about unitarily inequivalent solutions. These are used, for example, in attempting to describe differences on the two sides of a phase transition, such as occur in cosmological theories. In such applications the status of the relevant physical modality will be crudely analogous to that of the constant g in simpler classical theories incorporating Newton's law of gravitation with a slowly changing value of g or with alternative values for g that might arise in some other way. In such theories the value of g is not fixed of physical necessity in the same sense that the exponent of R is fixed of physical necessity. But at any one time the kind of contingency involved in the value of g in such a theory is radically different from and in particular more rigid than the kind of contingency involved in the fact that two gravitating particles might have had slightly different masses or physical separation. Analogous things are to be said for the unitarily inequivalent representations that are put to work in describing only a limited or "distant" kind of physical contingency as opposed to the more garden-variety physical contingency associated with the physical possibility of alternative states of the sort one seeks to influence in the laboratory. For those applications of quantum field theory that in this way support a qualified but still robust contrast between garden variety physical contingency and much more strict modalities there will continue be a robust sense in which the OVQF is a value of something like an L-determinable but itself corresponding to something like a P- (but not an L-) determinable.

What about unitarily equivalent solutions, all formally equally good but distinct solutions of the field equations satisfying the same commutation relations? Many unitary transformations can be seen as passive transformations, simply changing our representational conventions. These cases will all fall under the qualification, "once the convention fixing choices have been made." Such formally distinct solutions are no more distinct physical solutions than are distinct solutions arising from a change in units.

It would be very hard to argue in anything like complete generality that unitary transformations of the OVQF must always be seen as effecting no more than a transformation of one representation of a set of physical facts to a formally distinct representation of the same physical facts. In quantum theories unitary transformations are used for so many different things, including active transformations. But I trust that, even with out going into details, it will be plausible that in many extremely simple models unitary transformations of the OVQF will clearly all function to describe changes in conventions in the representation of the physical circumstances. In these very simple models, we can use the oversimplified recipe for interpreting the OVQF. One should then look at what is added or changed when one moves from such a very simple model to one or another more realistic one. Changes which have repercussions for how to think about the physical necessity or contingency of the OVQF will have corresponding repercussions for how to modify what one reads off the oversimplified picture. I see no reason to expect a uniform result from working out all of the very diverse applications of quantum field theory. However, I suspect that in a great many cases, while there will be refinements in the description of the relevant characterizations of physical necessity and contingency, the result will still look very like the oversimplified picture.

Here is an example. One can describe the S-matrix as a unitary transformation connecting the way the in-quantum field is transformed into the out-quantum field. Each of these fields are solutions to the free field equation, and the unitary transformation is interpreted actively, specifying how the incoming field is affected by the interaction. But, of course, this unitary transformation does not specify the transformation of in-states to out-states individually. It describes them collectively, giving at one stroke how any more specific in-coming state will be transformed into the resulting specific out-going state. In this case the in-field covers all physical possibilities for a free incoming state, the out-field describes all physical possibilities for a free outgoing state, and the unitary transformation from the one to the other describes, not some one transformation from some determinate, physically contingent in-state to the ensuing determinate out-state, but all the transformations for any determinate physically possible in- and out-states that are contingently possible once the interaction has been set.

8.3 The OVQF is Not an "Active Agent"

I have argued that once one gets the content of the OVQF as a field configuration clearly into view the analogy to one among many contingently possible solutions of classical field equations appears misleading, at least in a wide range of applications. In this section I want to explore in more detail what can and cannot plausibly be said about the OVQF in its status as a field configuration. For I have heard some of my colleagues[5] make strong claims for the nature of this configuration—that it counts as a field configuration in the same robust *causal* sense as does a determinate configuration of the electromagnetic field, that it is a state of the field which is to be thought of as operating causally, as an "active agent". I don't know of any such statement in print. But this attitude appears to have some currency; and even if few hold such an active agent reading, seeing in more detail what is problematic about such an attitude will sharpen our interpretive understanding of the ways in which the OVQF is like and unlike other things we call 'fields' and how these contrasts are related to other interpretive presuppositions.

In the last section I concluded that the OVQF has a status as a determinate field configuration only insofar as this configuration is seen as a determinate value of an L-field determinable; and this same configuration then also itself counts as a P-field determinable that covers the full range of physically possible determinate physical states. This conclusion should already make one cautious about any "active" agent reading: It's not facts about what is physically possible that makes things happen. If any "agent" is "acting" in any given case it would have to be the more specifically realized physical state. In this section I want in addition to present a number of other ways in which we can test this characterization. To do this I will need to make one more distinction, in the first instance concerning dispositions, which will then apply to characterize distinctions in the ways in which fields can be understood. In this section I will streamline exposition by speaking in terms of the oversimplified picture of the last section, before the qualifications.

8.3.1 *S- and C-Dispositions*

To attribute a disposition is, at least in part, to characterize relevant possibilities. To say that something is, for example, fragile or water soluble is to say that it is possible for the object to break or to dissolve in water. Some advocate a more robust view of dispositions according to which a disposi-

[5] I'll be discreet and not mention names....

tional property is really the specific physical configuration or collection of characteristics that, in the presence of triggering conditions, are causally responsible for the display of the disposition. For example, on this view the dispositional property of salt to which we refer when we say that salt is water-soluble is the complex of physical characteristics of salt that are causally responsible for it going into solution when placed in water. On this view, to attribute a disposition is to attribute the realizing physical characteristics even if we don't know what they are. We use the expression, "salt is water soluble", to refer to characteristics that are logically independent of, yet causally responsible for, salt going into solution when placed in water, in much the same way in which we use the expression "the cause of the fire" to refer to the logically independent conditions which are causally responsible for a fire.

Let's mark this contrast as follows:

S-dispositions ('S' for 'structural'): To attribute an S-disposition to something is to do no more than to characterize it in terms of a certain range of eventualities that are physically possible for the thing in various circumstances. To attribute an S- disposition to something is to specify no more than such a range of physical possibilities and not to attribute any specific concrete mechanism by which the possibilities are realized. The specified eventualities may be very simple or highly structured.

C-dispositions ('C' for 'Categorical'): A C-disposition is a characteristic that involves not only an S-disposition, but also some specific physical mechanism which, together with triggering conditions, brings about the display of the disposition. Mechanisms are here understood broadly to include any properties, complex of properties, or physical arrangements that, together with triggering conditions, constitute the causal agent responsible for the display of the disposition in question.

I should emphasize that as I want to use 'S-dispositions', attributing an S-disposition does not commit one to the existence (or to the non-existence) of any C-disposition, that is, to the existence (or non-existence) of any mechanism in terms of which the range or pattern of the S-disposition's eventualities might be understood.

In what follows it will be helpful for understanding the contrasting roles of the OVQF as a P-field-determinable and also as a field configuration of an L-field-determinable to sketch the ways in which S- and C-dispositions play different roles in explanations.

Citing a C-disposition can be explanatory in the sense of citing a cause—since a C-disposition is the mechanism which, together with triggering conditions, brings about the display of the disposition. In simple cases S-dispositions would appear to underwrite no more than pseudo- explanations, as exemplified in the familiar example of citing opium's "dormitive virtue" in an effort to explain why smoking opinion tends to put people to sleep.

This example is misleading. Appeal to the structure of a range of possibilities can be explanatory. Suppose that the possibilities covered in the opium case are structured: more opium, more drowsiness. Then even in the simple case in which opium is characterized as possessing the dormitive virtue one can explain why one person is drowsier than another by appealing to the fact that the first smoked more opium than the second. When a range of possibilities has more structure, the structure will support more robust explanations. For example, evolutionary accounts of the development of species don't work by giving causes—they don't cite the relevant physical mutations nor the environmental events which cause less fit organisms to die. They work by laying out a structure of differential possibilities which enable one to see how things are going to work out in the kind of way they do, whatever the underlying physical mechanisms. This is one way of seeing how the epithet "survival of the fittest" can function as explanatory even though one turns around and characterizes fitness in terms of survival rates. Thus understood, *fitness* is an S-disposition but nonetheless functions in illuminating structural explanations.[6,7]

[6]I am here urging a broader notion of "structural explanation" than that proposed by McMullin in his (1978). McMullin requires at least tacit appeal to the existence of a causal mechanism. I do not. (See also, Glennan (1996) in this regard.) There may, of course, be a substantive disagreement between us if McMullin insists that, contrary to what I have urged, without appealing, at least tacitly, to some underlying causal embodiment the appeal to structure fails to be explanatory.

[7]One more contrast between S- and C-dispositions, which however will not play any role in our further considerations, is that C-dispositions are generally individuated more finely than the corresponding S-dispositions. When the same range of possible occurrences arises through the operation of distinct physical mechanisms one will have more than one C-dispositions falling under the same S-dispostion. All this is, of course, also relative to principles for individuating distinct causal mechanisms.

8.3.2 *Minimalist and Active Agent Construals of Fields*

We are now ready to apply these considerations about dispositions to clarify differing ways of thinking about fields. The relevant distinctions are deeply rooted in the history of field concepts, a history which I am not competent to present. But historical accuracy is not needed to spell out the relevant distinctions which we now appear to have.[8]

The minimalist conception of fields can be thought of as a response to the discomfort, in the 17th and 18th century, with the action at a distance that seemed to be described by Newton's theory.[9] The field concept supplies at least a formal resolution to the worry. Application of Newton's laws to one mass formally supports a description in terms of a gravitational potential, characterized in terms of a field value at each spatial point. The values at the spatial points (or at least the values in arbitrarily small volumes) can in turn be thought of in terms of dispositions—they say what will happen to various objects if placed at the spatial points in question.

But the disposition is only an S-disposition! In the Newtonian framework there is no underlying mechanism at each spatial point (or in spatial volumes), not even in the broad sense in which I have been understanding the term, 'mechanism'. There are no independently characterizable properties of the spatial points which can be thought of in the spirit of an active causal agent. There is no finite propagation or, again in the Newtonian framework, associated mass/energy at each spatial point. Instead the Newtonian gravitation potential associated with one fixed mass is a restatement of the potential repercussions of Newton's laws in application to the given mass, potentially for any other mass which could be introduced at various spatial points. That is, the gravitational potential field provides a catalogue of possibilities delimited by Newton's laws and the given mass associated with the potential field in question.

The Newtonian gravitational potential field is a field in a minimalist sense —it is characterized by specification of something like S-dispositions

[8]I am drawing heavily on a short essay by Ernan McMullin (to appear) that served as introduction to the 1999 Seven Pines Symposium on Field Theory mentioned in the acknowledgement below. I am eager to acknowledge this source of my thinking on the present subject while wanting to be very careful not attribute any claims to McMullin . As I will be putting the relevant distinctions very much to my own use, readers must strictly refrain from making any inferences from my use of the ideas to McMullin's views on these issues.

[9]What McMullin calls the "gravity dilemma". See McMullin, (1989, passim.).

at the spatial points. That is, the Newtonian gravitational potential field gives the structure of the physically contingent possibilities which are set once one has fixed on the mass the gravitational "effects" of which are summarized by the gravitational potential. As such it need not be explanatorily barren—it can underwrite structural explanations. But it does not describe the causal action of some active agent at each spatial point. Instead it recapitulates certain consequences of Newton's laws in application to a source mass the "effects" of which are characterized in terms of the associated potential field.

On a very different active agent conception of field configurations, each space-time point is thought of as occupied by a substance, or as possessing a property, that has the power to cause various effects. The exemplar is the Eulerian conception of a velocity field of a physical, space-filling fluid.[10] In Euler's case it is the fluid, the stuff, which is pushing around the fluid at neighboring space-time points. In such a case the field properties do cause, or have the power to cause, the effects which we associate with the field. If the field is then characterized in such a case in terms of attribution of dispositions to space-time points, we have an example of an active agent conception of the field understood in terms of C-dispositions.

Let's illustrate further with the case of the electromagnetic field, which I suspect many think of as an active agent. On the one hand, the field can be characterized dispositionally, for example in terms of the dispositions to produce forces on test charges. On the other hand, one often thinks of the determinate state of the electromagnetic field at one point as the cause of motion of charged particles and as causally responsible for (as well as undergoing effects caused by) the state at neighboring points. These two ways of thinking about the electromagnetic field are expressed together if the dispositional characterization is given in terms of C-dispositions. This active agent conception is supported by (or perhaps in part consists in) the way we think about the propagation of electromagnetic effects at a finite velocity. Other characteristics which seem to go with, or to serve as marks of, the active agent conception are the facts that the field can be influenced locally, and only locally, that it carries associated mass/energy, and that, as in the

[10]For a statement and elaboration of this way of thinking about fields, and its contrast with what I am calling the minimalist conception, see for example, Hesse (1961, p. 192 ff).

example of the velocity field of a fluid, multiple properties are involved.[11] Such considerations incline many to think of a determinate configuration of the electromagnetic filed as active and involving C-dispositions in a way that a determinate configuration of the Newtonian gravitational field does not.

8.3.3 *The Character of the OVQF as a Field Configuration of an L-Field-Determinable*

What now of the idea that the OVQF bears interpreting as an active agent? The OVQF counts as a determinate field configuration only if taken as a determinate value of an L-field-determinable. The OVQF also counts as a full P-field-determinable, holding of physical necessity and covering all the physically possible more determinate physical states. But physical possibilities aren't causes. As a determinate field configuration of an L-field-determinable, the OVQF, in its guise as a determinate field configuration, should count, at best, as minimalist. The OVQF, as a configuration of an L-field determinable, can be characterized in terms of complex dispositions associated with the space-time points—and, most importantly, dispositional relations among the points. But these are S-, not C-dispositions. We certainly do not envisage any underlying mechanism. The OVQF does not physically propagate. The contribution of the OVQF to explanations is structural, not causal in any intuitive sense of "active causal power". The argument for this, in sum, is that the OVQF carries, if not exactly, then something quite close or similar to the information conveyed by the field equations as constrained by the canonical commutation relations, and such information contributes the basis for structural, not active agent casual explanations.

8.4 What Has and Has Not Been Shown

An awful lot of this discussion is skating on somewhat thin ice. I've offered nothing better than a very rough and ready, intuitive grip on the contrast

[11] Hesse (1961, pp. 203 ff) attributes this way of thinking to Faraday. Maxwell described the electromagnetic field in terms of analogical mechanical models or imaginary illustrations that involve states of motion and tension of a postulated ether. It is possible that our sense of the electromagnetic field as an active agent field is still colored by this history.

between active agent causes vs. explanations which work in terms of laying out the range or structure of what is physically possible and the associated distinction between minimalist vs. active agent conceptions of fields. This grip is based only on a few examples, not on any detailed analysis. Consequently, the thin ice will immediately give way under the weight of a rejection of any real distinction between causes cited in the spirit of appeal to mechanisms vs. the playing out of the structure of physical possibilities.

There are many who reject these distinctions, for example those with strong Humean sympathies. A Humean holds that causal connections are only constant conjunctions—facts about what regularly co-occurs with what, and on anything like such a conception of cause the needed distinction between structural and mechanistic explanations threatens to collapse.[12] Those who see no viable distinction between structural and causal explanations, between S- and C-dispositions, and so between minimalist and active agent conceptions of fields, will conclude that the question of whether or not the OVQF should be thought of as an "active agent" is a non-issue, in as much as they will insist that the alleged issue turns on an alleged distinction with no content.

In sum, though my analysis leaves much room for refinement, I take it to support the following conclusions: For those who recognize distinctions like that between logical and physical possibility, the OVQF is at once a P-field-determinable and a determinate field configuration of an L-field-determinable, where the P/L-determinable contrast will follow the corresponding modal distinctions. This conclusion turned on the supposition that the OVQF is an encoding of the facts covered by the field equations with the strictures imposed by the commutation relations, or on facts with similar modal status. Such facts themselves underwrite structural, not causal explanations. Consequently, those who recognize anything like the sort of distinction between minimalist and active agent conceptions of fields that I have tried to indicate should conclude that, in its guise as a field configuration of an L-field-determinable, the OVQF counts as minimalist.

[12]But see Glennan (1996) who argues that when the structure of constant conjunctions is taken to arise through the working of some physical mechanism a conscientious Humean can distinguish between brute constant conjunction laws which hold at a "bottom level" and structured causal connections associated with structured physical mechanisms. McMullin (1978) is here also relevant.

Acknowledgements

This paper owes a great deal to conversations with Gordon Fleming and Andrew Wayne. Jeremy Butterfield provided extensive comments. My thinking on this subject was also substantially guided by discussions at the 1999 Seven Pines Symposium. I am greatly indebted to all its participants and to the support for the symposium provided by its founder, Lee Gohlike.

References

Fleming, G. (2002): Comments on Paul Teller's book, "An Interpretive Introduction to Quantum Field Theory". This volume.

Hesse, M. (1961): *Forces and Fields*, Edinburgh: Nelson.

Glennan, S. S. (1996): Mechanisms and the nature of causation, *Erkenntnis*, 49-71.

McMullin, E. (1978): Structural explanation, *American Philosophical Quarterly*, 139-147

McMullin, E. (1989): The Explanation of Distant Action: Historical Notes. In J. Cushing and E. McMullin, eds., *Philosophical Consequences of Quantum Theory: Reflections on Bell's Theorem.* Notre Dame: University of Notre Dame Press.

McMullin, E. (to appear). *The historical origins of the field concept in physics.*

Wayne, Andrew (2002): A naive view of the quantum field. This volume.

PART 3

Relativity, Measurement and Renormalization

This page is intentionally left blank

Chapter 9

On the Nature of Measurement Records in Relativistic Quantum Field Theory

Jeffrey A. Barrett
University of California at Irvine

Abstract. *A resolution of the quantum measurement problem would require one to explain how it is that we end up with determinate records at the end of our measurements. Metaphysical commitments typically do real work in such an explanation. Indeed, one should not be satisfied with one's metaphysical commitments unless one can provide some account of determinate measurement records. I will explain some of the problems in getting determinate records in relativistic quantum field theory and pay particular attention to the relationship between the measurement problem and a generalized version of Malament's theorem.*

9.1 Introduction

Does relativistic quantum field theory tell us that the world is made of fields or particles or something else? One difficulty in answering this is that physical theories typically do not pin down a single preferred ontology. This can be seen in classical mechanics where we are some 350 years on, and we have nothing like a canonical metaphysics for the theory. Are the fundamental entities of classical mechanics point particles or are they extended objects? Does the theory tell us that there is an absolute substantival space or are positions only relative to other objects? Of course, part of the problem here is that it is not entirely clear what classical mechanics is. But even if one does the reconstruction work that it would take to get a sharp formal theory, one can always provide alternative metaphysical interpretations. This can be seen as an aspect of a general underdetermination problem: not only are physical theories typically underdetermined by empirical evidence, but one's ontological commitments are typically underdetermined by the physical theory one adopts.

If our physical theories are in fact always subject to interpretation, then one might take the debate over the proper ontology of relativistic quantum field theory to be futile. While there is something right in this reaction, metaphysical considerations have in the past proven important to understanding and to clearly formulating physical theories, and we could certainly use all the clarity we can get in finding a satisfactory formulation of relativistic quantum field theory. If one could cook up a satisfactory ontology for some formulation of relativistic quantum field theory, then it would mean that that formulation of the theory could be understood as descriptive of the physical world, and in the context of relativistic quantum field theory, this would be something new. What is required here is not just showing that the particular theory is logically consistent by providing a model; what we want is to show that the theory could be descriptive of *our* physical world.

One of the features of our world is that we have determinate measurement records. We perform experiments, record the results, then compare these results against the predictions of our physical theories. Measurement records then should somehow show up in the ontology that we associate with our best physical theory. Indeed, if not for the existence of such records, it would be difficult to account for the possibility of empirical science all.

I mention this aspect of our world because the existence of determinate records is something that is difficult to get in nonrelativistic quantum mechanics and more difficult to get in relativistic quantum mechanics. The problem of getting determinate measurement records is the quantum measurement problem.

Metaphysics typically does real work in solutions to the quantum measurement problem by providing the raw material for explaining how it is that we have determinate measurement records. We see this in solutions to the quantum measurement problem in nonrelativistic quantum mechanics. In Bohm's theory it is the always determinate particle positions that provide determinate measurement records. In many-world interpretations it is the determinate facts in the world inhabited by a particular observer that determines the content of that observer's records.

The point here is just that in quantum mechanics one's metaphysical commitments must be sensitive to how one goes about solving the measurement problem. Indeed, it seems to me that no metaphysics for relativistic quantum field theory can be considered satisfactory unless determinate measurement records somehow show up in one's description of the world. Put another way, one must have a solution to the quantum measurement problem before one can trust any specific interpretation of relativistic quantum field theory.

9.2 The Measurement Problem

The measurement problem arises in nonrelativistic quantum mechanics when one tries to explain how it is that we get determinate measurement records. If the deterministic unitary dynamics (the time-dependent Schrödinger equation in nonrelativistic quantum mechanics) described all physical interactions, then a measurement would typically result in an entangled superposition of one's measuring apparatus recording mutually contradictory outcomes. If one has a good measuring apparatus that starts ready to make a measurement, the linear dynamics predicts one would typically end up with something like:

$$\sum a_i |p_i\rangle_S \, |``p_i"\rangle_M \tag{9.1}$$

This is a state where (the measured system S having property p_1 and the measuring apparatus M recording that the measured system has property p_1) is superposed with (the measured system S having property p_2 and the measuring apparatus M recording that the measured system has property p_2) etc. And this clearly does not describe the measuring apparatus M as recording any particular determinate measurement record.[1]

This indeterminacy problem is solved on the standard von Neumann-Dirac formulation of nonrelativistic quantum mechanics by stipulating that the state of the measured system randomly collapses to an eigenstate of the observable being measured whenever one makes a measurement, where the probability of collapse to the state $|p_k\rangle_S \, |``p_k"\rangle_M$ is $|a_k|^2$. It is this collapse of the state that generates a determinate measurement record ($|p_k\rangle_S \, |``p_k"\rangle_M$ is a state

where S determinately has property p_k and M determinately records that S has property p_k). But it is notoriously difficult to provide an account of how and when collapses occur that does not look blatantly ad hoc and

[1] See Barrett (1999) for a detailed account of what it means to have a good measuring device and why it would necessarily end up in this sort of entangled state.

even harder to provide and account that is consistent with the demands of relativity.[2]

If there is no collapse of the quantum mechanical state on measurement, then one might try adding something to the usual quantum-mechanical state that represents the values of the determinate physical records. This so-called hidden variable would determine the value of one's determinate measurement record even when the usual quantum-mechanical state represents an entangled superposition of incompatible records. But it has proven difficult to describe the evolution of this extra component of the physical state in a way that is compatible with relativity.[3]

It is orthodox dogma that it is only possible to reconcile quantum mechanics and relativity in the context of a quantum field theory, where the fundamental entities are fields rather than particles.[4] While there may be other reasons for believing that we need a field theory in order to reconcile quantum mechanics and relativity (and we will consider one of these shortly), relativistic quantum field theory does nothing to solve the quantum measurement problem and it is easy to see why.

In relativistic quantum field theory one starts by adopting an appropriate relativistic generalization of the nonrelativistic unitary dynamics. The relativistic dynamics describes the relations that must hold between quantum-mechanical field states in neighboring space-time regions. By knowing how the field states in different space-time regions are related, one can then make statistical predictions concerning expected correlations between measurements performed on the various field quantities. But relativistic quantum field theory provides no account whatsoever for how determinate measurement records might be generated.

The problem here is analogous to the problem that arises in nonrelativistic quantum mechanics. If the possible determinate measurement records are supposed to be represented by the elements of some set of orthogonal field configurations, then there typically are no determinate measurement

[2]For two related attempts to get a collapse theory that satisfies the demands of relativity see Aharonov and Albert (1980) and Fleming (1988 and 1996).

[3]Much of the literature on this topic is concerned with either trying to find a version of Bohm's theory that is compatible with relativity or trying to explain why strict compatibility between the two theories is not really necessary. See Barrett (2000) for a discussion of Bohm's theory and relativity.

[4]This is the position expressed, for example, by Steven Weinberg (1987, 78–9). See also David Malament (1996, 1–9)

records since (given the relativistic unitary dynamics) the state of the field in a given space-time region will typically be an entangled superposition of different elements of the orthogonal set of field configurations. An appropriate collapse of the field would generate a determinate local field configuration which might in turn represent a determinate measurement result, but such an evolution of the state would violate the relativistic unitary dynamics. And, as it is usually presented, relativistic quantum field theory has nothing to say about the conditions under which a such a collapse might occur, nor does it have anything to say about how such an evolution might be made compatible with relativity. One might try adding a new physical parameter to the usual quantum mechanical state that represents the values of one's determinate measurement records. But relativistic quantum field theory has nothing to say about how to do this or about how one might then give a relativity-compatible dynamics for the new physical parameter.[5]

So relativistic quantum field theory does nothing to solve the quantum measurement problem. Indeed, because of the additional relativistic constraints, accounting for determinate measurement records is more difficult than ever.

In what follows, I will explain another sense in which the metaphysics of relativistic quantum mechanics must be sensitive to measurement considerations and why we are far from having a clear account of measurement in relativistic quantum mechanics.

9.3 Malament's Theorem

David Malament (1996) presented his local entities no-go theorem in defense of the dogma that a field ontology, not a particle ontology, is appropriate to relativistic quantum mechanics. The theorem follows from four apparently weak conditions that most physicists would expect to be satisfied by

[5]That one can predict statistical correlations between measurement results but cannot explain the determinate measurement results has led some (see Rovelli (1997) and Mermin (1998) for example), to conclude that relativistic quantum field theory (and quantum mechanics more generally) predicts statistical correlations without there being anything that is in fact statistically correlated—"correlations without correlata." The natural objection to this conclusion is that the very notion of there being statistical correlations between measurement records presumably requires that there be determinate measurement records.

the structure one would use to represent the state of a single particle in relativistic quantum mechanics. If these conditions are satisfied, then the theorem entails that the probability of finding the particle in any closed spatial region must be zero, and this presumably violates the assumption that there is a (detectable) particle at all. Malament thus concludes that a particle ontology is inappropriate for relativistic quantum mechanics.

A version of Malament's theorem can be proven that applies equally well to point particles or extended objects. I will describe this version of the theorem without proof.[6] The statement of the theorem below and its physical interpretation follows Malament (1996) with a few supporting comments.

Let M be Minkowski space-time, and let $\mathcal{H}$ be a Hilbert space where a ray in $\mathcal{H}$ represents the pure state of the object S. Let P_Δ be the projection operator on $\mathcal{H}$ that represents the proposition that the object S would be detected to be *entirely* within spatial set Δ if a detection experiment were performed. Relativistic quantum mechanics presumably requires one to satisfy at least the following four conditions.

(1) *Dynamics Translation Covariance Condition*: For all vectors a in M and for all spatial sets Δ

$$P_{\Delta+a} = U(a)P_\Delta U(-a) \tag{9.2}$$

where $a \mapsto U(a)$ is a strongly continuous, unitary representation in $\mathcal{H}$ of the translation group in M and $\Delta+a$ is the set that results from translating Δ by the vector a.

This condition stipulates that the dynamics is represented by a family of unitary operators. More specifically, it says that the projection operator that represents the proposition that the object would be detected within spatial region $\Delta + a$ can be obtained by a unitary transformation that depends only on a of the projection operator that represents the proposition that the object would be detected within region Δ. Note that if this condition is universally satisfied, then there can be no collapse of the quantum-mechanical state.

[6] The proof of this version of the theorem is essentially the same as the proof in Malament (1996). The only difference is the physical interpretation of P_Δ. Malament's theorem relies on a lemma by Borchers (1967).

(2) *Finite Energy Condition*: For all future-directed time-like vectors a in M, if $H(a)$ is the unique self-adjoint operator satisfying

$$U(ta) = \exp -itH(a), \tag{9.3}$$

then the spectrum of H(a) is bounded below.

$H(a)$ is the Hamiltonian of the system S. It represents the energy properties of the system and determines the unitary dynamics (by the relation above). Supposing that the spectrum of the Hamiltonian is bounded below amounts to supposing that S has a finite (energy) ground state.

(3) *Hyperplane Localizability Condition*: If Δ_1 and Δ_2 are disjoint spatial sets in the same hyperplane,

$$P_{\Delta_1}P_{\Delta_2} = P_{\Delta_2}P_{\Delta_1} = \mathbf{0} \tag{9.4}$$

where $\mathbf{0}$ is the zero operator on $\mathcal{H}$.

This condition is supposed to capture the intuition that a single object S cannot be entirely within any two disjoint regions at the same time (relative to any inertial frame). This is presumably part of what it would mean to say that there is *just one* spatially extended object.

(4) *General Locality Condition*: If Δ_1 and Δ_2 are any two disjoint spatial sets that are spacelike related (perhaps not on the same hyperplane!),

$$P_{\Delta_1}P_{\Delta_2} = P_{\Delta_2}P_{\Delta_1}. \tag{9.5}$$

Relativity together with what it means to be an object presumably requires that if an object were detected to be entirely within one spatial region, then since an object cannot travel faster than light, it could not also be detected to be entirely within a disjoint, space-like related region in any inertial frame. If this is right, then one would expect the following to hold

($\diamondsuit$) *Relativistic Object Condition*: For *any* two spacelike related spatial regions Δ_1 and Δ_2 (not just any two in the same hyperplane!)

$$P_{\Delta_1}P_{\Delta_2} = P_{\Delta_2}P_{\Delta_1} = \mathbf{0}. \tag{9.6}$$

Condition ($\diamondsuit$) is strictly stronger than the conjunction of conditions (3) and (4). The idea behind condition (4) is that even if it *were* possible to detect S to be entirely within two disjoint spacelike related spatial regions and if condition (3) were still satisfied (because the two detectors were in different inertial frames and Δ_1 and Δ_2 were consequently not in the same hyperplane), then the probability of detecting the object to be entirely within Δ_1 should at least be statistically independent of the probability of detecting it to be entirely within Δ_2. That is, proving the theorem from conditions (3) and (4) rather than the strictly stronger (but very plausible!) condition ($\Diamond$) allows for the possibility that particle detection in a particular space-time region might be hyperplane dependent. While this is certainly something that Malament would want to allow for (since he was responding to Fleming's hyperplane-dependent formulation of quantum mechanics), it is probably not a possibility that most physicists would worry about much. If this is right, then one might be perfectly happy replacing conditions (3) and (4) by condition ($\Diamond$).

The theorem is that if conditions (1)–(4) are satisfied (or conditions (1), (2), and ($\Diamond$)), then $P_\Delta = \mathbf{0}$ for all compact closed spatial sets Δ. This means that the only extended object possible (or, perhaps better, the only *detectable* extended object possible) is one with infinite extension. And this conclusion is taken to favor a field ontology. It may also have curious implications for the nature of one's measurement records in relativistic quantum mechanics. Or it may be that getting determinate measurement records in relativistic quantum mechanics requires one to violate one or more of the four conditions that make the theorem possible.

9.4 Measurement Records

In the broadest sense, a good measurement consists in correlating the state of a record with the physical property being measured. The goal is to produce a detectable, reliable, and stable record. It might be made in terms of ink marks on paper, the final position of the pointer on a measuring device, the bio-chemical state of an observer's brain, or the arrangement of megaliths on the Salisbury Plain; but whatever the medium, useful measurement records must be detectable (so that one can know the value of the record), reliable (so that one can correctly infer the value of the physical property that one wanted to measure), and stable (so that one can make reliable inferences concerning physical states at different times). Such measurement records provide the evidence on which empirical science is grounded.

Consider the following simple experiment where I test my one-handed typing skills. This experiment involves, as all do, making a measurement.

> The time it took me to type this sentence one-handed (because I am holding a stopwatch in the other hand) up to the following colon: **41.29 seconds**.

I am indeed a slow typist, but that is not the point. The point is that I measured then recorded how long it took me to type the above sentence fragment one-handed; and because I have a determinate, detectable, reliable, and stable record token, I know how long it took to type the sentence fragment, and you do too if you have interacted with the above token of the measurement record in an appropriate way.

Setting aside the question of exactly what it might mean for a measurement record to be reliable and stable, let's consider the detectability condition. For a record token to be detectable, it must presumably be the sort of thing one can find. And in order to be the sort of thing one can find, the presence or absence of a detectable record token R must presumably be something that can be represented in quantum mechanics as a projection operator on a finite spatial region. That is, there must be a projection operator R_Δ that represents the proposition that there is an R-record in region Δ. This is apparently just part of what it means for a record to be detectable in relativistic quantum mechanics.

Now consider the bold-faced typing-speed record token above. It is detectable. Not only can you find and read it, but you can find and read it in a finite time. If we rule out superluminal effects, then it seems that the *detected record token* must occupy a finite spatial region. Call this spatial region S. Given the way that observables are represented in relativistic quantum mechanics, this means that there must be a projection operator R_S that represents the proposition that there is a token of the **41.29 seconds** record in region S.

The problem with this is that Malament's theorem tells us that there can be no such record-detection operator. More specifically, it tells us that $R_\Delta = \mathbf{0}$ for all closed sets Δ, which means that the probability of finding the record token in the spatial set S is zero. Indeed, the probability of finding the (above!?) record token anywhere is zero. But how can this be if there is in fact a detectable record token? And if there is no detectable record token, then how can you and I know the result of my typing-speed measurement as we both presumably do?

A natural reaction would be to deny the assumption that a detectable record token is a detectable entity that occupies a finite spatial region and insist that in relativistic quantum field theory, as one would expect, all determinate record tokens are represented in the determinate configuration of some unbounded field. After all, this is presumably how records would have to be represented in *any* field theory.

More concretely, couldn't a determinate measurement record be represented, say, in the local configuration of an unbounded field? Sure, but there are a couple of problems one would still have to solve in order to have a satisfactory account of determinate measurement records.

One problem, of course, is the old one. Given the unitary dynamics and the standard interpretation of states, relativistic quantum field theory would typically not predict a determinate local field configuration in a spacetime region. But let's set the traditional measurement problem aside for a moment and suppose that we can somehow cook up a formulation of the theory where one typically does have a determinate local field configurations at the end of a measurement.

If one could somehow get determinate local field configurations that are appropriately correlated, then one could explain how it is possible for me to know my typing speed by stipulating that my mental state supervenes on the determinate value of some field quantity in a some spatial region region that, in turn, is reliably correlated with my typing speed. So not only is it possible for a local field configuration to represent a determinate measurement result, but one can explain how it is possible for an observer to know the value of the record by stipulating an appropriate supervenience relation between mental and physical states. What more could one want?

It seems to me that one should ultimately want to explain how our actual measurements might yield determinate records. But to do this, one needs an account of measurement records that makes sense of the experiments that we in fact perform. The problem here is that our measurement records seem to have locations; they are the sort of things that one can find, lose, and move from one place to another. Indeed, we use their spacio-temporal properties to individuate our records. In order to know how fast I typed the sentence, I must be able to find the right record, and this (apparently) amounts to looking for it in the right place. It seems then that we know where our records are, and this is good because, given the way that we individuate our records, one must know where a record is in order to read it and to know what one is reading! This is just a point about our experimental practice and conventions.

So it seems that our actual records are in fact detectable in particular spacetime regions. But if this is right, then there must be detection-of-a-record-at-a-location operators (R_Δ that represent the proposition that there is a record in region Δ). And if these are subject to Malament's theorem, then we have a puzzle: there apparently cannot be detectable records of just the sort that we take ourselves to have.

This is particularly puzzling when one considers the sort of records that are supposed to provide the empirical support for relativistic quantum field theory itself. These records are supposed to include such things as photographs of the trajectories of fundamental particles, but if there are no detectable spacio-temporal entities, then how could there be a photographic trajectory record with a detectable shape? The shape of the trajectory is supposed to represent all of the empirical evidence that one has, but it seems, at least at first pass, that there can be no detectable entities with determinate shapes given Malament's theorem. From this perspective, the problem is to account for our particular-like measurement phenomena using a theory that apparently has nothing with particle-like structure.

While Malament's theorem arguably does nothing to prohibit an entity from having a determinate position, it does seem to prohibit anything from having a *detectable* position. But detectable positions are just what our records apparently have: they are typically individuated by position, so one must be able to find a record at a location to read it and to know what one is reading, and, given our practice and conventions, the records themselves are typically supposed to be made in terms of the detectable position or shape of something.

One might argue that one does not need to know where a record is in order to set up the appropriate correlations needed to read a record or that one can know where the record is and thus set up the appropriate correlations to read it without the position of the record itself being detectable. And while one might easily see how each of these lines of argument would go, it seems to me that our actual practice ultimately renders such arguments implausible. If I forget what my typing speed was, then I need to find a stable reliable record, and, given the way that I recorded it and the way that I individuate my records, in order to find one, I must do a series of position detection observations: Only if I can find *where* the record token is, can I then determine *what* it is.

The situation is made more puzzling by the fact that we are used to treating observers themselves as localizable entities in order to get specific empirical predictions out of our physical theories.[7] The location an observer occupies provides the observer with the spacio-temporal perspective that we use to explain why the world appears the way it does to that observer and not the way it might to another. We also use the fact that an observer occupies a location to explain why her empirical knowledge has spatio-temporal constraints.[8]

If detectable spatio-temporal objects are incompatible with relativistic quantum mechanics, then the challenge is to explain why it seems that we and those physical objects to which we have the most direct epistemic access (our measurement records) are just such objects.[9] As far as I can tell, it is possible that all observers and their records are somehow represented in field configurations; it is just unclear how the making, finding, and reading of such records is supposed to work in relativistic quantum field theory. Perhaps one could argue that observers and their records have only approximate positions and that this is enough for us to individuate them (and make sense of what it means for theory to be empirically adequate for a given observer), then argue that there is nothing analogous to Malament's theorem in relativistic quantum mechanics that prevents there from being detectable entities with only approximately determinate positions. Our standard talk of detectable localized objects might then be translated into the physics of such quasi-detectable, quasi-localized objects. But again this would require some careful explaining.

On the other hand, it may well be that none of this matters after all. The difficult problem, the one on which the solution to the others must

[7]Consider, for example, Galileo comparing the motions of the planets against theoretical predictions. That the observer has a specifiable relative position is needed for the theory to make any empirical predictions, and without comparing such predictions against what he actually sees, he would never be able to judge the empirical merits of the theory.

[8]If *I* am represented in the configuration of an *unbounded* relativistic field, then why don't I know what is happening around α-Centauri right now (in my inertial frame—whatever that might be if I have no fully determinate position!)? After all, on this representation of me, I would be there now. Or, for that matter, why would I not know what will happen here two minutes from now?

[9]Note that the problem of explaining how we could have the records we have without there being detectable spatio-temporal objects is more basic than the problem of explaining why it appears that there are detectable particles or other extended objects since the only way that we know of other spatio-temporal objects is via our records of them (in terms of patches of photographic pigment, or patterns of neurons firing on one's retina, etc.).

hang, is the one we set aside earlier in this section. The real problem for finding a satisfactory interpretation of relativistic quantum field theory is the quantum measurement problem.

While theorems like Malament's might be relevant to what metaphysical morals one should draw from relativistic quantum mechanics, whether such theorems hold or not is itself contingent on how one goes about solving the quantum measurement problem. A collapse formulation of quantum mechanics would, for example, typically violate condition (1): The dynamics translation covariance condition is an assumption concerning how physical states in different space-time regions are related, and it is incompatible with a collapse of the quantum mechanical state on measurement. But if we might have to violate the apparently weak and obvious assumptions that go into proving Malament's theorem in order to get a satisfactory solution to the measurement problem, then all bets are off concerning the applicability of the theorem to the detectable entities that inhabit our world.[10]

The upshot of these reflections is that we are very nearly back where we started: one cannot trust any specific metaphysical conclusions one draws from relativistic quantum field theory without a solution to the quantum measurement problem, and we have every reason to suppose that the constraints imposed by relativity will make finding a satisfactory solution more difficult than ever.

9.5 Conclusion

An adequate resolution of the quantum measurement problem would explain how it is that we have the determinate measurement records that we take ourselves to have. It has proven difficult to find a satisfactory resolution of the measurement problem in the context of nonrelativistic quantum mechanics, and relativistic quantum mechanics does nothing to make the task any easier. Indeed, the constraints imposed by relativity make explaining how we end up with the determinate, detectable, physical records all the more difficult.

Since one's ontological commitments typically do real work in proposed resolutions to the quantum measurement problem in nonrelativistic quan-

[10]That a solution to the quantum measurement problem might require one to violate such conditions might be taken to illustrate how difficult it is to solve the measurement problem and satisfy relativistic constraints as they are typically understood.

tum mechanics, it would be a mistake to try to draw any conclusions concerning the proper ontology of relativistic quantum field theory without a particular resolution to the measurement problem in mind. This point is clearly made by the fact that one cannot even know whether the so-called local entities no-go theorems are relevant to one's theory if one does not know what to do about the quantum measurement problem.

Acknowledgements

I would like to thank David Malament and Rob Clifton for several discussions and many valuable comments on an earlier draft of this paper.

References

Aharonov, Y. and Albert, D. (1980): States and Observables in Relativistic Quantum Field Theory. *Physical Review* D 21:3316–24.

Aharonov, Y. and Albert, D. (1981): Can We Make Sense out of the Measurement Process in Reltivistic Quantum Mechanics. *Physical Review* D 24:359–70.

Barrett, J. A. (1996): Empirical Adequacy and the Availability of Reliable Records in Quantum Mechanics. *Philosophy of Science*, 63:49–64.

Barrett, J. A. (1999): *The Quantum Mechanics of Minds and Worlds.* Oxford: Oxford University Press.

Barrett, J. A. (2000): The Persistence of Memory: Surreal Trajectories in Bohm's theory. *Philosophy of Science*, 67: 680-703.

Borchers, H. J. (1967): A Remark on Theorem by B. Misra. *Communications in Mathematical Physics* 4:315–323.

Dickson, M. and R. Clifton (1998): Lorentz-Invariance in Modal Interpretations. In D. Dieks and P. Vermaas (eds.), *The Modal Interpretation of Quantum Mechanics*, Dordrecht: Kluwer, 9–48.

Fleming, G. (1988): Hyperplane-Dependent Quantized Fields and Lorentz Invariance. In H. R. Brown and R. Harré (eds.), *Philosophical Foundations of Quantum Field Theory*, Oxford: Clarendon Press, 93–115.

Malament, D. B. (1996): In Defense of Dogma: Why There Cannot be a Relativistic Quantum Mechanics of (Localizable) Particles. In R. Clifton (ed.), *Perspectives on Quantum Reality*, The Netherlands: Kluwer Academic Publishers, 1–10.

Mermin, D. (1998): What is Quantum Mechanics Trying to Tell Us? arxiv.org preprint quant-ph/9801057 v2 (2 Sep 1998).

Rovelli, C. (1997): Relational Quantum Mechanics. arxiv.org preprint quant-ph/9609002 v2 (24 Feb 1997).

Weinberg, S. (1987): *Elementary Particles and the Laws of Physics, The 1986 Dirac Memorial Lectures*, Cambridge: Cambridge University Press.

This page is intentionally left blank

Chapter 10

No Place for Particles in Relativistic Quantum Theories?

Hans Halvorson
Princeton University

Rob Clifton
University of Pittsburgh

Abstract. *David Malament (1996) has recently argued that there can be no relativistic quantum theory of (localizable) particles. We consider and rebut several objections that have been made against the soundness of Malament's argument. We then consider some further objections that might be made against the generality of Malament's conclusion, and we supply three no-go theorems to counter these objections. Finally, we dispel potential worries about the counterintuitive nature of these results by showing that relativistic quantum field theory itself explains the appearance of "particle detections."*

10.1 Introduction

It is a widespread belief, at least within the physics community, that there is no relativistic quantum theory of (localizable) particles; and, thus, that the only relativistic quantum theory is a theory of fields. This belief has received much support in recent years in the form of rigorous no-go theorems by Malament (1996) and Hegerfeldt (1998a, 1998b). In particular, Hegerfeldt shows that in a generic quantum theory (relativistic or nonrelativistic), if there are states with localized particles, and if there is a lower bound on the system's energy, then superluminal spreading of the wavefunction must occur. Similarly, Malament shows the inconsistency of a

[0]The editors thank the University of Chicago Press for permission to reprint this article from *Philosophy of Science*, 69(1), 2002.

few intuitive desiderata for a relativistic quantum mechanics of (localizable) particles. Thus, it appears that quantum theory engenders a fundamental conflict between relativistic causality and localizability.

What is the philosophical lesson of this conflict between relativistic causality and localizability? On the one hand, if we believe that the assumptions of Malament's theorem must hold for any empirically adequate theory, then it follows that our world cannot be correctly described by a particle theory. On the other hand, if we believe that our world *can* be correctly described by a particle theory, then one (or more) of the Malament's assumptions must be false. Malament clearly endorses the first response; that is, he argues that his theorem entails that there is no relativistic quantum mechanics of localizable particles (insofar as any relativistic theory precludes act-outcome correlations at spacelike separation). Others, however, have argued that the assumptions of Malament's theorem need not hold for any relativistic, quantum-mechanical theory (cf. Fleming and Butterfield 1999), or that we cannot judge the truth of the assumptions until we resolve the interpretive issues of elementary quantum mechanics (cf. Barrett 2001).

We do not think that these objections to the soundness of Malament's argument are cogent. However, there are other tacit assumptions of Malament's theorem that some might be tempted to question. For example, Malament's theorem depends on the assumption that there is no preferred inertial reference frame, which some believe to have very little empirical support (cf. Cushing 1996). Furthermore, Malament's theorem establishes only that there is no relativistic quantum mechanics in which particles can be completely localized in spatial regions with sharp boundaries; it leaves open the possibility that there might be a relativistic quantum mechanics of "unsharply" localized particles.

In this paper, we present two new no-go theorems which show that these tacit assumptions of Malament's theorem are not needed to sustain an argument against localizable particles. First, we derive a no-go theorem against localizable particles that does not assume the equivalence of all inertial frames (Theorem 3). Second, we derive a no-go theorem that shows that there is no relativistic quantum mechanics of unsharply localized particles (Theorem 5).

However, it would be a mistake to think that these results show — or, are intended to show — that a field ontology, rather than a particle ontology, is appropriate for relativistic quantum theories. While these results show that there is no position observable that satisfies relativistic constraints, quantum field theories — both relativistic *and* non-relativistic — already reject the notion of a position observable in favor of localized field observables. Thus, our first two results have nothing to say about the possibility that relativistic quantum field theory (RQFT) might permit a "particle interpretation," in which localized particles are supervenient on the underlying localized field observables. To exclude this latter possibility, we formulate (in Section 10.6) a necessary condition for a quantum field theory to permit a particle interpretation, and we then show that this condition fails in any relativistic theory (Theorem 6).

Presumably, any empirically adequate theory must be able to reproduce the predictions of special relativity and of quantum mechanics. Therefore, our no-go results show that the existence of localizable particles is, strictly speaking, ruled out by the empirical data. However, in Section 10.7 we defuse this counterintuitive consequence by showing that RQFT itself explains how the illusion of localizable particles can arise, and how "particle talk" — although strictly fictional — can still be useful.

10.2 Malament's Theorem

Malament's theorem shows the inconsistency of a few intuitive desiderata for a relativistic quantum mechanics of (localizable) particles. It strengthens previous results (e.g., Schlieder 1971) by showing that the assumption of "no superluminal wavepacket spreading" can be replaced by the weaker assumption of "microcausality," and by making it clear that Lorentz invariance is not needed to derive a conflict between relativistic causality and localizability.

In order to present Malament's result, we assume that our background spacetime M is an affine space, with a foliation $\mathcal{S}$ into spatial hyperplanes. This will permit us to consider a wide range of relativistic (e.g., Minkowski) as well as non-relativistic (e.g., Galilean) spacetimes. The pure states of our quantum-mechanical system are given by rays in some Hilbert space $\mathcal{H}$. We assume that there is a mapping $\Delta \mapsto E_\Delta$ of bounded subsets of hyperplanes in M into projections on $\mathcal{H}$. We think of E_Δ as representing the proposition that the particle is localized in Δ; or, from a more operational point of view, E_Δ represents the proposition that a position measurement is certain to find the particle within Δ. We also assume that there is a strongly continuous representation $\mathbf{a} \mapsto U(\mathbf{a})$ of the translation group of M in the unitary operators on $\mathcal{H}$. Here strong continuity means that for any unit vector $\psi \in \mathcal{H}$, $\langle \psi, U(\mathbf{a})\psi \rangle \to 1$ as $\mathbf{a} \to 0$; and it is equivalent (via Stone's theorem) to the assumption that there are energy and momentum observables for the particle. If all of the preceding conditions hold, we say that the triple $(\mathcal{H}, \Delta \mapsto E_\Delta, \mathbf{a} \mapsto U(\mathbf{a}))$ is a *localization system* over M.

The following conditions should hold for any localization system — either relativistic or non-relativistic — that describes a single particle.

Localizability: If Δ and Δ' are disjoint subsets of a single hyperplane, then $E_\Delta E_{\Delta'} = 0$.

Translation covariance: For any Δ and for any translation $\mathbf{a}$ of M, $U(\mathbf{a})E_\Delta U(\mathbf{a})^* = E_{\Delta+\mathbf{a}}$.

Energy bounded below: For any timelike translation $\mathbf{a}$ of M, the generator $H(\mathbf{a})$ of the one-parameter group $\{U(t\mathbf{a}) : t \in \mathbb{R}\}$ has spectrum bounded from below.

We recall briefly the motivation for each of these conditions. "Localizability" says that the particle cannot be detected in two disjoint spatial sets at a given time. "Translation covariance" gives us a connection between the symmetries of the spacetime M and the symmetries of the quantum-mechanical system. In particular, if we displace the particle by a spatial translation $\mathbf{a}$, then the original wavefunction ψ will transform to some wavefunction $\psi_{\mathbf{a}}$. Since the statistics for a displaced detection experiment should be identical to the original statistics, we have $\langle \psi, E_\Delta \psi \rangle = \langle \psi_{\mathbf{a}}, E_{\Delta+\mathbf{a}} \psi_{\mathbf{a}} \rangle$. By Wigner's theorem, however, the symmetry is implemented by some unitary operator $U(\mathbf{a})$. Thus, $U(\mathbf{a})\psi = \psi_{\mathbf{a}}$, and $U(\mathbf{a})E_\Delta U(\mathbf{a})^* = E_{\Delta+\mathbf{a}}$. In the case of time translations, the covariance condition entails that the particle has unitary dynamics. (This might seem to beg the question against a collapse interpretation of quantum mechanics; we dispell this worry at the end of this section.) Finally, the "energy bounded below" condition asserts that, relative to any inertial observer, the particle has a lowest possible energy state. If it were to fail, we could extract an arbitrarily large amount of energy from the particle as it drops down through lower and lower states of energy.

We now turn to the "specifically relativistic" assumptions needed for Malament's theorem. The special theory of relativity entails that there is a finite upper bound on the speed at which (detectable) physical disturbances can propagate through space. Thus, if Δ and Δ' are distant regions of space, then there is a positive lower bound on the amount of time it should take for a particle localized in Δ to travel to Δ'. We can formulate this requirement precisely by saying that for any timelike translation $\mathbf{a}$, there is an $\epsilon > 0$ such that, for every state ψ, if $\langle \psi, E_\Delta \psi \rangle = 1$ then $\langle \psi, E_{\Delta' + t\mathbf{a}} \psi \rangle = 0$ whenever $0 \leq t < \epsilon$. This is equivalent to the following assumption.

Strong causality: If Δ and Δ' are disjoint subsets of a single hyperplane, and if the distance between Δ and Δ' is nonzero, then for any timelike translation $\mathbf{a}$, there is an $\epsilon > 0$ such that $E_\Delta E_{\Delta' + t\mathbf{a}} = 0$ whenever $0 \leq t < \epsilon$.

(Note that strong causality entails localizability.) Although strong causality is a reasonable condition for relativistic theories, Malament's theorem requires only the following weaker assumption (which he himself calls "locality").

Microcausality: If Δ and Δ' are disjoint subsets of a single hyperplane, and if the distance between Δ and Δ' is nonzero, then for any timelike translation $\mathbf{a}$, there is an $\epsilon > 0$ such that $[E_\Delta, E_{\Delta' + t\mathbf{a}}] = 0$ whenever $0 \leq t < \epsilon$.

If E_Δ can be measured within Δ, microcausality is equivalent to the assumption that a particle detection measurement within Δ cannot influence the statistics of particle detection measurements performed in regions that are spacelike to Δ (see Malament 1996, 5). Thus, a failure of microcausality would entail the possibility of act-outcome correlations at spacelike separation. Note that both strong causality and microcausality make sense for non-relativistic spacetimes (as well as for relativistic spacetimes); though, of course, we should not expect either causality condition to hold in the non-relativistic case.

Theorem 1 ***(Malament).*** *Let $(\mathcal{H}, \Delta \mapsto E_\Delta, \mathbf{a} \mapsto U(\mathbf{a}))$ be a localization system over Minkowski spacetime that satisfies:*

(1) Localizability
(2) Translation covariance
(3) Energy bounded below
(4) Microcausality

Then $E_\Delta = 0$ for all Δ.

Thus, in every state, there is no chance that the particle will be detected in any local region of space. As Malament claims, this serves as a *reductio ad absurdum* of any relativistic quantum mechanics of a single (localizable) particle.

10.2.1 *The Soundness of Malament's Argument*

Several authors have claimed that the assumptions of Malament's theorem need not hold for any relativistic, quantum-mechanical theory of particles. For example, Dickson (1997) argues that a 'quantum' theory does not need a position operator (equivalently, a system of localizing projections) in order to treat position as a physical quantity; Barrett (2001) argues that translation covariance is suspect; and Fleming and Butterfield (1999) argue that the microcausality assumption is not warranted by special relativity. We now show, however, that none of these arguments is decisive against the assumptions of Malament's theorem.

Dickson (1997, 214) cites the Bohmian interpretation of the Dirac equation as a counterexample to the claim that any 'quantum' theory must represent position by an operator. In order to see what Dickson might mean by this, recall that the Dirac equation admits both positive and negative energy solutions. If $\mathcal{H}$ denotes the Hilbert space of *all* (both positive and negative energy) solutions, then there is a self-adjoint position operator Q on $\mathcal{H}$ defined by $Q\psi(\mathbf{x}) = \mathbf{x} \cdot \psi(\mathbf{x})$ (cf. Thaller 1992, 7). If, however, we restrict to the Hilbert space $\mathcal{H}_{\text{pos}} \subset \mathcal{H}$ of positive energy solutions, then the probability density given by the Dirac wavefunction does not correspond to a self-adjoint position operator (Thaller 1992, 32).

According to Holland (1993, 502), the lack of a position operator on $\mathcal{H}_{\text{pos}}$ precludes a Bohmian interpretation of $\psi(\mathbf{x})$ as a probability amplitude for finding the particle in an elementary volume $d^3\mathbf{x}$ around $\mathbf{x}$. Rather, the Bohmian approach makes use of the position observable Q on the full Hilbert space $\mathcal{H}$ of both positive and negative energy solutions. Thus, it appears that Dickson was simply mistaken to claim that the Bohmian interpretation of the Dirac equation dispenses with a position observable. Furthermore, since the Bohmian interpretation of the Dirac equation violates the energy bounded below condition, it does not provide a counterexample to Malament's theorem.

However, Dickson could have developed his argument by appealing to the positive energy subspace $\mathcal{H}_{\text{pos}}$. In this case, we *can* talk about particle positions despite the fact that we do not have a position observable in the usual sense. In particular, we will show in Section 10.5 that, for talk about positions, it suffices to have a family of "unsharp" localization observables. (And, yet, we shall show that relativistic quantum theories do not permit even this attenuated notion of localization.)

Barrett (2001) argues that the significance of Malament's theorem cannot be assessed until we have solved the measurement problem:

> If we might have to violate the apparently weak and obvious assumptions that go into proving Malament's theorem in order to get a satisfactory solution to the measurement problem, then all bets are off concerning the applicability of the theorem to the detectable entities that inhabit our world. (Barrett 2001, 16)

In particular, a solution to the measurement problem may require that we abandon unitary dynamics. But if we abandon unitary dynamics, then translation covariance does not hold.

Unfortunately, it is not clear that we could avoid the upshot of Malament's theorem by moving to a collapse theory. Existing (non-relativistic) collapse theories take the empirical predictions of quantum theory seriously. That is, the "statistical algorithm" of quantum mechanics is assumed to be at least approximately correct; and collapse is introduced only to ensure that we obtain determinate properties at the end of a measurement. However, in the present case, Malament's theorem shows that any quantum theory predicts that if there are local particle detections, then act-outcome correlations are possible at spacelike separation. Thus, if a collapse theory is to reproduce these predictions, it too would face a conflict between localizability and relativistic causality.

Perhaps, then, Barrett is suggesting that the price of accommodating localizable particles might be a complete abandonment of unitary dynamics, *even at the level of a single particle.* In other words, we may be forced to adopt a collapse theory *without* having any underlying (unitary) quantum theory. But even if this is correct, it wouldn't count against Malament's theorem, which was intended to show that there is no relativistic *quantum* theory of localizable particles. Furthermore, noting that Malament's theorem requires unitary dynamics is one thing; it would be quite another thing to provide a model in which there *are* localizable particles — at the price of non-unitary dynamics — but which is also capable of reproducing the well-confirmed quantum interference effects at the micro-level. Until we have such a model, pinning our hopes for localizable particles on a failure of unitary dynamics is little more than wishful thinking.

Like Barrett, Fleming (Fleming and Butterfield 1999, 158ff) disagrees with the reasonableness of Malament's assumptions. Unlike Barrett, however, Fleming provides a concrete model in which there are localizable particles (viz., using the Newton-Wigner position operator as a localizing observable) and in which the microcausality assumption fails. Nonetheless, Fleming argues — contra Malament — that this failure of microcausality is perfectly consistent with relativistic causality.

According to Fleming, the property "localized in Δ" (represented by E_Δ) need not be detectable within Δ. As a result, $[E_\Delta, E_{\Delta'}] \neq 0$ does not entail that it is possible to send a signal from Δ to Δ'. However, by claiming that local *beables* need not be local *observables*, Fleming undercuts the primary utility of the notion of localization, which is to indicate those physical quantities that are operationally accessible in a given region of spacetime. Indeed, it is not clear what motivation there could be — aside from indicating what is locally measurable — for assigning observables to spatial regions. If E_Δ is *not* measurable in Δ, then why should we say that "E_Δ is localized in Δ"? Why not say instead that "E_Δ is localized in Δ'" (where $\Delta' \neq \Delta$)? Does either statement have any empirical consequences and, if so, how do their empirical consequences differ? Until these questions are answered, we maintain that local beables are always local observables; and a failure of microcausality *would* entail the possibility of act-outcome correlations at spacelike separation. (For a more detailed argument along these lines, see Halvorson 2001, Section 6.)

10.2.2 *Tacit Assumptions of Malament's Theorem*

The objections to the four assumptions of Malament's theorem are unconvincing. By any reasonable understanding of special relativity and of quantum theory, these assumptions should hold for any theory that is capable of reproducing the predictions of both theories. Nonetheless, we anticipate that further objections could be directed against the more or less tacit assumptions of Malament's theorem.

As we noted earlier, Malament's theorem does not make use of the full structure of Minkowski spacetime (e.g., Lorentz invariance). However, the following example shows that the theorem fails if there is a preferred inertial reference frame.

Example 10.1 Let $M = \mathbb{R}^1 \oplus \mathbb{R}^3$ be full Newtonian spacetime with a distinguished timelike direction **a**. To any set of the form $\{(t, x) : x \in \Delta\}$, with $t \in \mathbb{R}$, and Δ a bounded open subset of $\mathbb{R}^3$, we assign the spectral projection E_Δ of the position operator for a particle in three dimensions. Thus, the conclusion of Malament's theorem is false, while both the microcausality and localizability conditions hold. Let $P_0 = 0$, and for $i = 1, 2, 3$, let $P_i = -i(d/dx_i)$. For any four-vector **b**, let $U(\mathbf{b}) = \exp\{i(\mathbf{b} \cdot \mathbf{P})\}$, where

$$\mathbf{b} \cdot \mathbf{P} \; = \; b_0 P_0 + b_1 P_1 + b_2 P_2 + b_3 P_3. \tag{10.1}$$

Thus, translation covariance holds, and since the energy is identically zero, the energy condition trivially holds. (Note, however, that if M is *not* regarded as having a distinguished timelike direction, then this example violates the energy condition.) □

A brief inspection of Malament's proof shows that the following assumption on the affine space M is sufficient for his theorem to go through.

No absolute velocity: Let **a** be a spacelike translation of M. Then there is a pair $(\mathbf{b}, \mathbf{c})$ of timelike translations of M such that $\mathbf{a} = \mathbf{b} - \mathbf{c}$.

Despite the fact that "no absolute velocity" is a feature of both Galilean and Minkowski spacetimes, there are some who claim that the existence of a (undetectable) preferred reference frame is perfectly consistent with all current empirical evidence (cf. Cushing 1996). What is more, the existence of a preferred frame is an absolutely essential feature of a number of "realistic" interpretations of quantum theory (cf. Maudlin 1994, Chap. 7). Thus, this tacit assumption of Malament's theorem could be a source of contention for those wishing to maintain the possibility of a relativistic quantum mechanics of localizable particles.

Second, some might wonder whether Malament's result is an artifact of special relativity, and whether a notion of localizable particles might be restored in the context of general relativity. Indeed, it is not difficult to see that Malament's result does *not* automatically generalize to arbitrary relativistic spacetimes.

To see this, suppose that M is an arbitrary globally hyperbolic manifold. (That is, M is a manifold that permits at least one foliation S into spacelike hypersurfaces). Although M will not typically have a translation group, we assume that M has a transitive Lie group G of diffeomorphisms. (Just as a manifold is locally isomorphic to $\mathbb{R}^n$, a Lie group is locally isomorphic to a group of translations.) We require that G has a representation $g \mapsto U(g)$ in the unitary operators on $\mathcal{H}$; and, the translation covariance condition now says that $E_{g(\Delta)} = U(g)E_\Delta U(g)^*$ for all $g \in G$. The following example then shows that Malament's theorem fails even for the very simple case where M is a two-dimensional cylinder.

Example 10.2 *Let $M = \mathbb{R} \oplus S^1$, where S^1 is the one-dimensional unit circle, and let G denote the Lie group of timelike translations and rotations of M.* It is not difficult to construct a unitary representation of G that satisfies the energy bounded below condition. (We can use the Hilbert space of square-integrable functions from S^1 into $\mathbb{C}$, and the procedure for constructing the unitary representation is directly analogous to the case of a single particle moving on a line.) Fix a spacelike hypersurface Σ, and let μ denote the normalized rotation-invariant measure on Σ. For each open subset Δ of Σ, let $E_\Delta = I$ if $\mu(\Delta) \geq 2/3$, and let $E_\Delta = 0$ if $\mu(\Delta) < 2/3$. Then localizability holds, since for any pair (Δ, Δ') of disjoint open subsets of Σ, either $\mu(\Delta) < 2/3$ or $\mu(\Delta') < 2/3$. □

Obviously, Examples 10.1 and 10.2 are not physically interesting counterexamples to Malament's theorem. In particular, in Example 10.1 the energy is identically zero, and therefore the probability for finding the particle in a given region of space remains constant over time. Similarly, in Example 10.2 the particle is localized in every region of space with volume greater than 2/3, and the particle is never localized in a region of space with volume less than 2/3. In the following two sections, then, we will formulate explicit conditions to rule out such pathologies, and we will use these conditions to derive a no-go theorem that applies to generic spacetimes.

10.3 Hegerfeldt's Theorem

Hegerfeldt's (1998a, 1998b) recent results on localization apply to arbitrary (globally hyperbolic) spacetimes, and they do not make use of the "no absolute velocity" condition. Thus, we will suppose henceforth that M is a globally hyperbolic spacetime, and we will fix a foliation $\mathcal{S}$ of M, as well as a unique isomorphism between any two hypersurfaces in this foliation. If $\Sigma \in \mathcal{S}$, we will write $\Sigma + t$ for the hypersurface that results from "moving Σ forward in time by t units"; and if Δ is a subset of Σ, we will use $\Delta + t$

to denote the corresponding subset of $\Sigma + t$. We assume that there is a representation $t \mapsto U_t$ of the time-translation group $\mathbb{R}$ in the unitary operators on $\mathcal{H}$; and we say that the localization system $(\mathcal{H}, \Delta \mapsto E_\Delta, t \mapsto U_t)$ satisfies *time-translation covariance* just in case $U_t E_\Delta U_{-t} = E_{\Delta+t}$ for all Δ and all $t \in \mathbb{R}$.

Hegerfeldt's result is based on the following root lemma.

Lemma 10.1 ***(Hegerfeldt).*** *Suppose that $U_t = e^{itH}$, where H is a self-adjoint operator with spectrum bounded from below. Let A be a positive operator (e.g., a projection operator). Then for any unit vector ψ, either*

$$\langle U_t\psi, AU_t\psi\rangle \neq 0\,, \qquad \text{for almost all } t \in \mathbb{R},$$

or

$$\langle U_t\psi, AU_t\psi\rangle = 0\,, \qquad \text{for all } t \in \mathbb{R}\,.$$

Hegerfeldt claims that this lemma has the following consequence for localization:

> If there exist particle states which are strictly localized in some finite region at $t = 0$ and later move towards infinity, then finite propagation speed cannot hold for localization of particles. (Hegerfeldt 1998a, 243)

Hegerfeldt's argument for this conclusion is as follows:

> Now, if the particle or system is strictly localized in Δ at $t = 0$ it is, a fortiori, also strictly localized in any larger region Δ' containing Δ. If the boundaries of Δ' and Δ have a finite distance and *if finite propagation speed holds* then the probability to find the system in Δ' must also be 1 for sufficiently small times, e.g. $0 \leq t < \epsilon$. But then [Lemma 10.1], with $A \equiv I - E_{\Delta'}$, states that the system stays in Δ' for *all* times. Now, we can make Δ' smaller and let it approach Δ. Thus we conclude that if a particle or system is at time $t = 0$ strictly localized in a region Δ, then finite propagation speed implies that it stays in Δ for all times and therefore prohibits motion to infinity. (Hegerfeldt 1998a, 242–243; notation adapted, but italics in original)

Let us attempt now to formalize this argument.

First, Hegerfeldt claims that the following is a consequence of "finite propagation speed": If $\Delta \subseteq \Delta'$, and if the boundaries of Δ and Δ' have a finite distance, then a state initially localized in Δ will continue to be localized in Δ' for some finite amount of time. We can capture this precisely by means of the following condition.

No instantaneous wavepacket spreading (NIWS): If $\Delta \subseteq \Delta'$, and the boundaries of Δ and Δ' have a finite distance, then there is an $\epsilon > 0$ such that $E_\Delta \leq E_{\Delta'+t}$ whenever $0 \leq t < \epsilon$.

(Note that NIWS plus localizability entails strong causality.) In the argument, Hegerfeldt also assumes that if a particle is localized in every one of a family of sets that "approaches" Δ, then it is localized in Δ. We can capture this assumption in the following condition.

Monotonicity: If $\{\Delta_n : n \in \mathbb{N}\}$ is a downward nested family of subsets of Σ such that $\bigcap_n \Delta_n = \Delta$, then $\bigwedge_n E_{\Delta_n} = E_\Delta$.

Using this assumption, Hegerfeldt argues that if NIWS holds, and if a particle is initially localized in some finite region Δ, then it will remain in Δ for all subsequent times. In other words, if $E_\Delta \psi = \psi$, then $E_\Delta U_t \psi = U_t \psi$ for all $t \geq 0$. We can now translate this into the following formal no-go theorem.

Theorem 2 ***(Hegerfeldt).*** *Suppose that the localization system $(\mathcal{H}, \Delta \mapsto E_\Delta, t \mapsto U_t)$ satisfies:*

(1) Monotonicity
(2) Time-translation covariance
(3) Energy bounded below
(4) No instantaneous wavepacket spreading

Then $U_t E_\Delta U_{-t} = E_\Delta$ for all $\Delta \subset \Sigma$ and all $t \in \mathbb{R}$.

(The proof of this and all subsequent theorems can be found in the appendix.)

Thus, conditions 1–4 can be satisfied only if the particle has trivial dynamics. The following Lemma then shows how to derive Malament's conclusion from Hegerfeldt's theorem.

Lemma 10.2 *Let M be an affine space. Suppose that the localization system $(\mathcal{H}, \Delta \mapsto E_\Delta, \mathbf{a} \mapsto U(\mathbf{a}))$ satisfies localizability, time-translation* *covariance, and no absolute velocity. For any bounded spatial set* Δ, *if* $U(\mathbf{a})E_\Delta U(\mathbf{a})^* = E_\Delta$ *for all timelike translations* $\mathbf{a}$ *of* M, *then* $E_\Delta = 0$.

Thus, if we add "no absolute velocity" to the assumptions of Hegerfeldt's theorem, then it follows that $E_\Delta = 0$ for all bounded Δ. However, NIWS is a stronger causality assumption than microcausality. In fact, while NIWS plus localizability entails strong causality (and hence microcausality), the following example shows that NIWS is not entailed by the conjunction of strong causality, monotonicity, time-translation covariance, and energy bounded below.

Example 10.3 Let Q, P denote the standard position and momentum operators on $\mathcal{H} = L_2(\mathbb{R})$, and let $H = P^2/2m$ for some $m > 0$. Let $\Delta \mapsto E^Q_\Delta$ denote the spectral measure for Q. Fix some bounded subset Δ_0 of $\mathbb{R}$, and let $E_\Delta = E^Q_\Delta \otimes E^Q_{\Delta_0}$ (a projection operator on $\mathcal{H} \otimes \mathcal{H}$) for all Borel subsets Δ of $\mathbb{R}$. Thus, $\Delta \mapsto E_\Delta$ is a (non-normalized) projection-valued measure. Let $U_t = I \otimes e^{itH}$, and let $E_{\Delta+t} = U_t E_\Delta U_{-t}$ for all $t \in \mathbb{R}$. It is clear that monotonicity, time-translation covariance, and energy bounded below hold. To see that strong causality holds, let Δ and Δ' be disjoint subsets of a single hyperplane Σ. Then,

$$E_\Delta U_t E_{\Delta'} U_{-t} \;=\; E^Q_\Delta E^Q_{\Delta'} \otimes E^Q_{\Delta_0} E^Q_{\Delta_0+t} \;=\; 0 \otimes E^Q_\Delta E^Q_{\Delta_0+t} \;=\; 0\,, \qquad (10.2)$$

for all $t \in \mathbb{R}$. On the other hand, $U_t E_\Delta U_{-t} \neq E_\Delta$ for any nonempty Δ and for any $t \neq 0$. Thus, it follows from Hegerfeldt's theorem that NIWS fails. □

Thus, we could not recapture the full strength of Malament's theorem simply by adding "no absolute velocity" to the conditions of Hegerfeldt's theorem.

10.4 Doing without "No Absolute Velocity"

Example 10.3 shows that Hegerfeldt's theorem fails if NIWS is replaced by strong causality (or by microcausality). On the other hand, Example 10.3 is hardly a physically interesting counterexample to a strengthened version of Hegerfeldt's theorem. In particular, if Σ is a fixed spatial hypersurface, and if $\{\Delta_n : n \in \mathbb{N}\}$ is a covering of Σ by bounded sets (i.e., $\bigcup_n \Delta_n = \Sigma$), then $\bigvee_n E_{\Delta_n} = I \otimes E_{\Delta_0} \neq I \otimes I$. Thus, it is not certain that the particle will be detected *somewhere or other* in space. In fact, if $\{\Delta_n : n \in \mathbb{N}\}$ is a covering of Σ and $\{\Pi_n : n \in \mathbb{N}\}$ is a covering of $\Sigma + t$, then

$$\bigvee_{n\in\mathbb{N}} E_{\Delta_n} = I \otimes E_{\Delta_0} \neq I \otimes E_{\Delta_0+t} = \bigvee_{n\in\mathbb{N}} E_{\Pi_n}. \tag{10.3}$$

Thus, the total probability for finding the particle somewhere or other in space can change over time.

It would be completely reasonable to require that $\bigvee_n E_{\Delta_n} = I$ whenever $\{\Delta_n : n \in \mathbb{N}\}$ is a covering of Σ. This would be the case, for example, if the mapping $\Delta \mapsto E_\Delta$ (restricted to subsets of Σ) were the spectral measure of some position operator. However, we propose that — at the very least — any physically interesting model should satisfy the following weaker condition.

Probability conservation: If $\{\Delta_n : n \in \mathbb{N}\}$ is a covering of Σ, and $\{\Pi_n : n \in \mathbb{N}\}$ is a covering of $\Sigma + t$, then $\bigvee_n E_{\Delta_n} = \bigvee_n E_{\Pi_n}$.

Probability conservation guarantees that there is a well-defined total probability for finding the particle somewhere or other in space, and this probability remains constant over time. In particular, if both $\{\Delta_n : n \in \mathbb{N}\}$ and $\{\Pi_n : n \in \mathbb{N}\}$ consist of pairwise disjoint sets, then the localizability condition entails that $\bigvee_n E_{\Delta_n} = \sum_n E_{\Delta_n}$ and $\bigvee_n E_{\Pi_n} = \sum_n E_{\Pi_n}$. In this case, probability conservation is equivalent to

$$\sum_{n\in\mathbb{N}} \text{Prob}^{\psi}(E_{\Delta_n}) = \sum_{n\in\mathbb{N}} \text{Prob}^{\psi}(E_{\Pi_n}), \tag{10.4}$$

for any state ψ. Note, finally, that it is reasonable to require probability conservation for both relativistic and non-relativistic models.[1] With this in mind, we can now formulate a no-go result that generalizes aspects of both Malament's and Hegerfeldt's theorems.

Theorem 3 *Suppose that the localization system* $(\mathcal{H}, \Delta \mapsto E_\Delta, t \mapsto U_t)$ *satisfies:*

(1) Localizability

[1] Probability conservation would fail if a particle could escape to infinity in a finite amount of time (cf. Earman 1986, 33). However, a particle can escape to infinity only if there is an infinite potential well, and this would violate the energy condition. Thus, given the energy condition, probability conservation should also hold for non-relativistic particle theories.

(2) Probability conservation
(3) Time-translation covariance
(4) Energy bounded below
(5) Microcausality

Then $U_t E_\Delta U_{-t} = E_\Delta$ *for all* Δ *and all* $t \in \mathbb{R}$.

If M is an affine space, and if we add "no absolute velocity" as a sixth condition in this theorem, then it follows that $E_\Delta = 0$ for all Δ (see Lemma 10.2). Thus, modulo the probability conservation condition, Theorem 3 recaptures the full strength of Malament's theorem. Moreover, we can now trace the difficulties with localization to microcausality alone: There are localizable particles only if it is possible to have act-outcome correlations at spacelike separation.

We now give examples to show that all five assumptions of Theorem 3 are essential for the result. (Example 10.1 shows that these assumptions can be simultaneously satisfied.) For simplicity, suppose that M is two-dimensional. (All examples work in the four-dimensional case as well.) Let Q, P be the standard position and momentum operators on $L_2(\mathbb{R})$, and let $H = P^2/2m$. Let Σ be a spatial hypersurface in M, and suppose that a coordinatization of Σ has been fixed, so that there is a natural association between each bounded open subset Δ of Σ and a corresponding spectral projection E_Δ of Q.

(1+2+3+4) (a) Consider the standard localization system for a single non-relativistic particle. That is, let $\Delta \mapsto E_\Delta$ (with domain the Borel subsets of Σ) be the spectral measure for Q. For $\Sigma + t$, set $E_{\Delta+t} = U_t E_\Delta U_{-t}$, where $U_t = e^{itH}$. (b) The Newton-Wigner approach to relativistic QM uses the standard localization system for a non-relativistic particle, only replacing the non-relativistic Hamiltonian $P^2/2m$ with the relativistic Hamiltonian $(P^2+m^2I)^{1/2}$, whose spectrum is also bounded from below.

(1+2+3+5) (a) For a mathematically simple (but physically uninteresting) example, take the first example above and replace the Hamiltonian $P^2/2m$ with P. In this case, microcausality trivially holds, since $U_t E_\Delta U_{-t}$ is just a shifted spectral projection of Q. (b) For a physically interesting example, consider the relativistic quantum theory of a single spin-1/2 electron (see Section 10.2). Due to the negative energy solutions of the Dirac equation, the spectrum of the Hamiltonian is not bounded from below.

(1+2+4+5) Consider the standard localization system for a nonrelativistic particle, but set $E_{\Delta+t} = E_{\Delta}$ for all $t \in \mathbb{R}$. Thus, we escape the conclusion of trivial dynamics, but only by disconnecting the (nontrivial) unitary dynamics from the (trivial) association of projections with spatial regions.

(1+3+4+5) (a) Let Δ_0 be some bounded open subset of Σ, and let E_{Δ_0} be the corresponding spectral projection of Q. When $\Delta \neq \Delta_0$, let $E_{\Delta} = 0$. Let $U_t = e^{itH}$, and let $E_{\Delta+t} = U_t E_{\Delta} U_{-t}$ for all Δ. This example is physically uninteresting, since the particle cannot be localized in any region besides Δ_0, including proper supersets of Δ_0. (b) See Example 10.3.

(2+3+4+5) Let Δ_0 be some bounded open subset of Σ, and let E_{Δ_0} be the corresponding spectral projection of Q. When $\Delta \neq \Delta_0$, let $E_{\Delta} = I$. Let $U_t = e^{itH}$, and let $E_{\Delta+t} = U_t E_{\Delta} U_{-t}$ for all Δ. Thus, the particle is always localized in every region other than Δ_0, and is sometimes localized in Δ_0 as well.

10.5 Are there Unsharply Localizable Particles?

We have argued that attempts to undermine the four explicit assumptions of Malament's theorem are unsuccessful. We have also now shown that the "no absolute velocity" condition is not necessary to rule out localizable particles. However, there is one further question that might arise concerning the soundness of Malament's argument. In particular, some might argue that it is possible to have a quantum-mechanical particle theory in the absence of a family $\{E_{\Delta}\}$ of localizing projections. What is more, one might argue that localizing projections represent an unphysical idealization — viz., that a "particle" can be completely contained in a finite region of space with a sharp boundary, when in fact it would require an infinite amount of energy to prepare a particle in such a state. Thus, there remains a possibility that relativistic causality can be reconciled with "unsharp" localizability.

To see how we can define "particle talk" without having projection operators, consider again the relativistic theory of a single spin-1/2 electron (where we now restrict to the subspace $\mathcal{H}_{\text{pos}}$ of positive energy solutions of the Dirac equation). In order to treat the argument '$\mathbf{x}$' of the Dirac wavefunction as an observable, it would be sufficient to define a probability amplitude and density for the particle to be found at $\mathbf{x}$; and these can be obtained from the Dirac wavefunction itself. That is, for a subset Δ of Σ, we set

$$\text{Prob}^{\psi}(\mathbf{x} \in \Delta) = \int_{\Delta} |\psi(\mathbf{x})|^2 d\mathbf{x}\,. \tag{10.5}$$

Now let $\Delta \mapsto E_{\Delta}$ be the spectral measure for the standard position operator on the Hilbert space $\mathcal{H}$ (which includes both positive and negative energy solutions). That is, E_{Δ} multiplies a wavefunction by the characteristic function of Δ. Let F denote the orthogonal projection of $\mathcal{H}$ onto $\mathcal{H}_{\text{pos}}$. Then,

$$\int_{\Delta} |\psi(\mathbf{x})|^2 d\mathbf{x} = \langle \psi, E_{\Delta}\psi \rangle = \langle \psi, FE_{\Delta}\psi \rangle, \tag{10.6}$$

for any $\psi \in \mathcal{H}_{\text{pos}}$. Thus, we can apply the standard recipe to the operator FE_{Δ} (defined on $\mathcal{H}_{\text{pos}}$) to compute the probability that the particle will be found within Δ. However, FE_{Δ} is *not* a projection operator. (In fact, it can be shown that FE_{Δ} does not have any eigenvectors with eigenvalue 1.) Thus, we do not need a family of *projection* operators in order to define probabilities for localization.

Now, in general, to define the probability that a particle will be found in Δ, we need only assume that there is an operator A_{Δ} such that $\langle \psi, A_{\Delta}\psi \rangle \in [0,1]$ for any unit vector ψ. Such operators are called *effects*, and include the projection operators as a proper subclass. Thus, we say that the triple $(\mathcal{H}, \Delta \mapsto A_{\Delta}, \mathbf{a} \mapsto U(\mathbf{a}))$ is an *unsharp localization system* over M just in case $\Delta \mapsto A_{\Delta}$ is a mapping from subsets of hyperplanes in M to effects on $\mathcal{H}$, and $\mathbf{a} \mapsto U(\mathbf{a})$ is a continuous representation of the translation group of M in unitary operators on $\mathcal{H}$. (We assume again for the present that M is an affine space.)

Most of the conditions from the previous sections can be applied, with minor changes, to unsharp localization systems. In particular, since the energy bounded below condition refers only to the unitary representation, it can be carried over intact; and translation covariance also generalizes straightforwardly. However, we will need to take more care with microcausality and with localizability.

If E and F are projection operators, $[E, F] = 0$ just in case for any state, the statistics of a measurement of F are not affected by a non-selective measurement of E and vice versa (cf. Malament 1996, 5). This fact, along with the assumption that E_Δ is measurable in Δ, motivates the microcausality assumption. For the case of an association of arbitrary effects with spatial regions, Busch (1999, Prop. 2) has shown that $[A_\Delta, A_{\Delta'}] = 0$ just in case for any state, the statistics for a measurement of A_Δ are not affected by a non-selective measurement of $A_{\Delta'}$ and vice versa. Thus, we may carry over the microcausality assumption intact, again seen as enforcing a prohibition against act-outcome correlations at spacelike separation.

The localizability condition is motivated by the idea that a particle cannot be simultaneously localized (with certainty) in two disjoint regions of space. In other words, if Δ and Δ' are disjoint subsets of a single hyperplane, then $\langle\psi, E_\Delta\psi\rangle = 1$ entails that $\langle\psi, E_{\Delta'}\psi\rangle = 0$. It is not difficult to see that this last condition is equivalent to the assumption that $E_\Delta + E_{\Delta'} \leq I$. That is,

$$\langle\psi, (E_\Delta + E_{\Delta'})\psi\rangle \leq \langle\psi, I\psi\rangle\,, \tag{10.7}$$

for any unit vector ψ. Now, it is an accidental feature of projection operators (as opposed to arbitrary effects) that $E_\Delta + E_{\Delta'} \leq I$ is equivalent to $E_\Delta E_{\Delta'} = 0$. Thus, the appropriate generalization of localizability to unsharp localization systems is the following condition.

Localizability: If Δ and Δ' are disjoint subsets of a single hyperplane, then $A_\Delta + A_{\Delta'} \leq I$.

That is, the probability for finding the particle in Δ, plus the probability for finding the particle in some disjoint region Δ', never totals more than 1. It would, in fact, be reasonable to require a slightly stronger condition, viz., the probability of finding a particle in Δ plus the probability of finding a particle in Δ' equals the probability of finding a particle in $\Delta \cup \Delta'$. If this is true for all states ψ, we have:

Additivity: If Δ and Δ' are disjoint subsets of a single hyperplane, then $A_\Delta + A_{\Delta'} = A_{\Delta\cup\Delta'}$.

With just these mild constraints, Busch (1999) was able to derive the following no-go result.

Theorem 4 ***(Busch).*** *Suppose that the unsharp localization system* $(\mathcal{H}, \Delta \mapsto A_\Delta, \mathbf{a} \mapsto U(\mathbf{a}))$ *satisfies localizability, translation covariance, energy bounded below, microcausality, and no absolute velocity. Then, for any* Δ, A_Δ *has no eigenvector with eigenvalue* 1.

Thus, it is not possible for a particle to be localized with certainty in any bounded region Δ. Busch's theorem, however, leaves it open whether there are (nontrivial) "strongly unsharp" relativistic localization systems. The following result shows that there are not.

Theorem 5 *Suppose that the unsharp localization system* $(\mathcal{H}, \Delta \mapsto A_\Delta, \mathbf{a} \mapsto U(\mathbf{a}))$ *satisfies:*

(1) Additivity
(2) Translation covariance
(3) Energy bounded below
(4) Microcausality
(5) No absolute velocity

Then $A_\Delta = 0$ *for all* Δ.

Theorem 5 shows that invoking the notion of unsharp localization does nothing to resolve the tension between relativistic causality and localizability. For example, we can now conclude that the positive energy Dirac theory violates microcausality.[2]

Unfortunately, Theorem 5 does not generalize to arbitrary globally hyperbolic spacetimes, as the following example shows.

Example 10.4 Let M be the cylinder spacetime from Example 10.2. Let G denote the group of timelike translations and rotations of M, and let $g \mapsto U(g)$ be a positive energy representation of G in the unitary operators on a Hilbert space $\mathcal{H}$. For any $\Sigma \in S$, let μ denote the normalized rotation-invariant measure on Σ, and let $A_\Delta = \mu(\Delta)I$. Then, conditions 1–5 of Theorem 5 are satisfied, but the conclusion of the theorem is false. □

The previous counterexample can be excluded if we require there to be a fixed positive constant δ such that, for each Δ, there is a state ψ with $\langle\psi, A_\Delta\psi\rangle \geq \delta$. In fact, with this condition added, Theorem 5 holds for any globally hyperbolic spacetime. (The proof is an easy modification of the proof we give in the Appendix.) However, it is not clear what physical motivation there could be for requiring this further condition. Note also that Example 10.4 has trivial dynamics; i.e., $U_t A_\Delta U_{-t} = A_\Delta$ for all Δ. We

[2] For any unit vector $\psi \in \mathcal{H}_{\text{pos}}$, there is a bounded set Δ such that $\int_\Delta |\psi|^2 d\mathbf{x} \neq 0$, and therefore $A_\Delta \neq 0$. On the other hand, additivity, translation covariance, energy bounded below, and no absolute velocity hold. Thus, it follows from Theorem 5 that microcausality fails.

conjecture that every counterexample to a generalized version of Theorem 5 will have trivial dynamics.

Theorem 5 strongly supports the conclusion that there is no relativistic quantum mechanics of a single (localizable) particle; and therefore that special relativity and quantum mechanics can be reconciled only in the context of a quantum field theory. However, neither Theorem 3 nor Theorem 5 says anything about the ontology of relativistic quantum field theories; they leave it fully open that such theories might permit an ontology of localizable particles. To eliminate this latter possibility, we will now proceed to present a more general result which shows that there are no localizable particles in *any* relativistic quantum theory.

10.6 Are there Localizable Particles in RQFT?

The localizability assumption is motivated by the idea that a "particle" cannot be detected in two disjoint spatial regions at once. However, in the case of a many-particle system, it is certainly possible for there to be particles in disjoint spatial regions. Thus, the localizability condition does not apply to many-particle systems; and Theorems 3 and 5 cannot be used to rule out a relativistic quantum mechanics of $n > 1$ localizable particles.

Still, one might argue that we could use E_Δ to represent the proposition that a measurement is certain to find that all n particles lie within Δ, in which case localizability should hold. Note, however, that when we alter the interpretation of the localization operators $\{E_\Delta\}$, we must alter our interpretation of the conclusion. In particular, the conclusion now shows only that it is not possible for all n particles to be localized in a bounded region of space. This leaves open the possibility that there are localizable particles, but that they are governed by some sort of "exclusion principle" that prohibits them all from clustering in a bounded spacetime region.

Furthermore, Theorems 3 and 5 only show that it is impossible to define *position operators* that obey appropriate relativistic constraints. But it does not immediately follow from this that we lack any notion of localization in relativistic quantum theories. Indeed,

> ...a position operator is inconsistent with relativity. This compels us to find another way of modeling localization of events. In field theory, we model localization by making the

> observables dependent on position in spacetime. (Ticiatti 1999, 11)

However, it is not a peculiar feature of RQFT that it lacks a position operator: All quantum field theories (both relativistic and non-relativistic) model localization by making the observables dependent on position in spacetime. Moreover, in the case of non-relativistic QFT, these "localized" observables suffice to provide us with a concept of localizable particles. In particular, for each spatial region Δ, there is a "number operator" N_Δ whose eigenvalues give the number of particles within the region Δ. Thus, we have no difficultly in talking about the particle content in a given region of space, despite the absence of any position operator.

Abstractly, a number operator N on $\mathcal{H}$ is any operator with eigenvalues $\{0, 1, 2, \dots\}$. In order to describe the number of particles locally, we require an association $\Delta \mapsto N_\Delta$ of subsets of spatial hyperplanes in M to number operators on $\mathcal{H}$, where N_Δ represents the number of particles in the spatial region Δ. If $\mathbf{a} \mapsto U(\mathbf{a})$ is a unitary representation of the translation group, we say that the triple $(\mathcal{H}, \Delta \mapsto N_\Delta, \mathbf{a} \mapsto U(\mathbf{a}))$ is a *system of local number operators* over M.

Note that a localization system $(\mathcal{H}, \Delta \mapsto E_\Delta, \mathbf{a} \mapsto U(\mathbf{a}))$ is a special case of a system of local number operators where the eigenvalues of each N_Δ are restricted to $\{0, 1\}$. Furthermore, if we loosen our assumption that number operators have a discrete spectrum, and instead require only that they have spectrum contained in $[0, \infty)$, then we can also include unsharp localization systems within the general category of systems of local number operators. Thus, a system of local number operators is the *minimal* requirement for a concept of localizable particles in any quantum theory.

In addition to the natural analogues of the energy bounded below condition, translation covariance, and microcausality, we will be interested in the following two requirements on a system of local number operators:[3]

[3]Due to the unboundedness of number operators, we would need to take some care in giving technically correct versions of the following conditions. In particular, the additivity condition should technically include the clause that N_Δ and $N_{\Delta'}$ have a common dense domain, and the operator $N_{\Delta \cup \Delta'}$ should be thought of as the self-adjoint closure of $N_\Delta + N_{\Delta'}$. In the number conservation condition, the sum $N = \sum_n N_{\Delta_n}$ can be made rigorous by exploiting the correspondence between self-adjoint operators and "quadratic forms" on $\mathcal{H}$. In particular, we can think of N as deriving from the upper bound of quadratic forms corresponding to finite sums.

Additivity: If Δ and Δ' are disjoint subsets of a single hyperplane, then $N_\Delta + N_{\Delta'} = N_{\Delta\cup\Delta'}$.

Number conservation: If $\{\Delta_n : n \in \mathbb{N}\}$ is a disjoint covering of Σ, then the sum $\sum_n N_{\Delta_n}$ converges to a densely defined, self-adjoint operator N on $\mathcal{H}$ (independent of the chosen covering), and $U(\mathbf{a})NU(\mathbf{a})^* = N$ for any timelike translation $\mathbf{a}$ of M.

Additivity asserts that, when Δ and Δ' are disjoint, the expectation value (in any state) for the number of particles in $\Delta \cup \Delta'$ is the sum of the expectations for the number of particles in Δ and the number of particles in Δ'. In the pure case, it asserts that the number of particles in $\Delta \cup \Delta'$ is the sum of the number of particles in Δ and the number of particles in Δ'. The "number conservation" condition tells us that there is a well-defined global number operator, and that its expectation values remains constant over time. This condition holds for any non-interacting model of QFT.

It is a well-known consequence of the Reeh-Schlieder theorem that relativistic quantum field theories do not admit systems of local number operators (cf. Redhead 1995). We will now derive the same conclusion from strictly weaker assumptions. In particular, microcausality is the only specifically relativistic assumption needed for this result. The relativistic spectrum condition — which requires that the spectrum of the four-momentum lie in the forward light cone, and which is used in the proof of the Reeh-Schlieder theorem — plays no role in our proof.[4]

Theorem 6 *Suppose that the system* $(\mathcal{H}, \Delta \mapsto N_\Delta, \mathbf{a} \mapsto U(\mathbf{a}))$ *of local number operators satisfies:*

(1) Additivity
(2) Translation covariance
(3) Energy bounded below
(4) Number conservation
(5) Microcausality
(6) No absolute velocity

Then $N_\Delta = 0$ *for all* Δ.

[4]Microcausality is not only sufficient, but also necessary for the proof that there are no local number operators. The Reeh-Schlieder theorem entails the cyclicity of the vacuum state. But the cyclicity of the vacuum state alone does not entail that there are no local number operators; we must also assume microcausality (cf. Halvorson 2001, Requardt 1982).

Thus, in every state, there are no particles in any local region. This serves as a *reductio ad absurdum* for a notion of localizable particles in any relativistic quantum theory.

Unfortunately, Theorem 6 is not the strongest result we could hope for, since "number conservation" holds only in the (trivial) case of non-interacting fields. However, we would need a more general approach in order to deal with interacting relativistic quantum fields, because (due to Haag's theorem; cf. Streater and Wightman 2000, 163) their dynamics are not unitarily implementable on a fixed Hilbert space. On the other hand, this hardly indicates a limitation on the generality of our conclusion, since Haag's theorem also entails that interacting models of RQFT have no number operators — not even a global number operator.[5] Still, it would be interesting to recover this conclusion (perhaps working in a more general algebraic setting) without using the full strength of Haag's assumptions.

10.7 Particle Talk without Particle Ontology

The results of the previous sections show that relativistic quantum theories do not admit (localizable) particles into their ontology. We also considered and rejected several objections to our characterization of relativistic quantum theories. Thus, we have yet to find a good reason to reject one of the premises of our argument against localizable particles. However, according to Segal (1964) and Barrett (2001), there are independent grounds for believing that there are localizable particles — and therefore for rejecting one of the premises of the no-go results.

The argument for localizable particles appears to be very simple: Our experience shows us that objects (particles) occupy finite regions of space. But the reply to this argument is just as simple: These experiences are illusory! Although no object is strictly localized in a bounded region of space, an object can be well-enough localized to give the appearance to us (finite observers) that it is strictly localized. In fact, RQFT *itself* shows how the "illusion" of localizable particles can arise, and how talk about localizable particles can be a useful fiction.

[5] If a total number operator exists in a representation of the canonical commutation relations, then that representation is quasiequivalent to a free-field (Fock) representation (Chaiken 1968). However, Haag's theorem entails that in relativistic theories, representations with nontrivial interactions are *not* quasiequivalent to a free-field representation.

In order to assess the possibility of "approximately localized" objects in RQFT, we shall now pursue the investigation in the framework of algebraic quantum field theory.[6] Here, one assumes that there is a correspondence $O \mapsto \mathcal{R}(O)$ between bounded open subsets of M and subalgebras of observables on some Hilbert space $\mathcal{H}$. Observables in $\mathcal{R}(O)$ are considered to be "localized" (i.e., measurable) in O. Thus, if O and O' are spacelike separated, we require that $[A, B] = 0$ for any $A \in \mathcal{R}(O)$ and $B \in \mathcal{R}(O')$. Furthermore, we assume that there is a continuous representation $\mathbf{a} \mapsto U(\mathbf{a})$ of the translation group of M in unitary operators on $\mathcal{H}$, and that there is a unique "vacuum" state $\Omega \in \mathcal{H}$ such that $U(\mathbf{a})\Omega = \Omega$ for all $\mathbf{a}$. This latter condition entails that the vacuum appears the same to all observers, and that it is the unique state of lowest energy.

In this context, a particle detector can be represented by an effect C such that $\langle \Omega, C\Omega \rangle = 0$. That is, C should register no particles in the vacuum state. However, the Reeh-Schlieder theorem entails that no positive local observable can have zero expectation value in the vacuum state. Thus, it is impossible to detect particles by means of local measurements; instead, we will have to think of particle detections as "approximately local" measurements.

If we think of an observable as representing a measurement procedure (or, more precisely, an equivalence class of measurement procedures), then the norm distance $\|C - C'\|$ between two observables gives a quantitative measure of the physical similarity between the corresponding procedures. (In particular, if $\|C - C'\| < \delta$, then the expectation values of C and C' never differ by more than δ.)[7] Moreover, in the case of real-world measurements, the existence of measurement errors and environmental noise make it impossible for us to determine precisely which measurement procedure we have performed. Thus, practically speaking, we can at best determine a neighborhood of observables corresponding to a concrete measurement procedure.

In the case of present interest, what we actually measure is always a local observable — i.e., an element of $\mathcal{R}(O)$, where O is bounded. However,

[6]For general information on algebraic quantum field theory, see (Haag 1992) and (Buchholz 2000). For specific information on particle detectors and "almost local" observables, see Chapter 6 of (Haag 1992) and Section 4 of (Buchholz 2000).

[7]Recall that $\|C - C'\|$ is defined as the supremum of $\|(C - C')\psi\|$ as ψ runs through the unit vectors in $\mathcal{H}$. It follows then from the Cauchy-Schwarz inequality that $|\langle \psi, (C - C')\psi \rangle| \leq \|C - C'\|$ for any unit vector ψ.

given a fixed error bound δ, if an observable C is within norm distance δ from some local observable $C' \in \mathcal{R}(O)$, then a measurement of C' will be practically indistinguishable from a measurement of C. Thus, if we let

$$\mathcal{R}_\delta(O) = \{C : \exists C' \in \mathcal{R}(O) \text{ such that } \|C - C'\| < \delta\}, \tag{10.8}$$

denote the family of observables "almost localized" in O, then 'FAPP' (i.e., 'for all practical purposes') we can locally measure any observable from $\mathcal{R}_\delta(O)$. That is, measurement of an element from $\mathcal{R}_\delta(O)$ can be simulated to a high degree of accuracy by local measurement of an element from $\mathcal{R}(O)$. However, for any local region O, and for any $\delta > 0$, $\mathcal{R}_\delta(O)$ *does* contain (nontrivial) effects that annihilate the vacuum.[8] Thus, particle detections can always be simulated by purely local measurements; and we can explain the appearance of macroscopically localized objects without assuming that there are localizable particles in the strict sense.

However, it may not be easy to pacify Segal and Barrett with a FAPP solution to the problem of localization. Both appear to think that the absence of localizable particles is not simply contrary to our manifest experience, but would undermine the very possibility of objective empirical science. For example, Segal claims that,

> ...it is an elementary fact, *without which experimentation of the usual sort would not be possible*, that particles are indeed localized in space at a given time. (Segal 1965, 145; our italics)

Furthermore, "particles would not be observable without their localization in space at a particular time" (1964, 139). In other words, experimentation involves observations of particles, and these observations can occur only if particles are localized in space. Unfortunately, Segal does not give any argument for these claims. It seems to us, however, that the moral we should draw from the no-go theorems is that Segal's account of observation is false. In particular, we do not observe particles; rather, there are 'observation events', and these observation events are consistent (to a good degree of accuracy) with the supposition that they are brought about by localizable particles.

[8]Suppose that $A \in \mathcal{R}(O)$, and let $A(\mathbf{x}) = U(\mathbf{x})AU(\mathbf{x})^*$. If f is a test function on M whose Fourier transform is supported in the complement of the forward light cone, then $L = \int f(\mathbf{x})A(\mathbf{x})d\mathbf{x}$ is almost localized in O and $\langle\Omega, L\Omega\rangle = 0$ (cf. Buchholz 2000, 7).

Like Segal, Barrett (2001) claims that we will have trouble explaining how empirical science can work if there are no localizable particles. In particular, Barrett claims that empirical science requires that we be able to keep an account of our measurement results so that we can compare these results with the predictions of our theories. Furthermore, we identify measurement records by means of their location in space. Thus, if there were no localized objects, then there would be no identifiable measurement records, and "...it would be difficult to account for the possibility of empirical science at all" (Barrett 2001, 3).

However, it's not clear what the difficulty here is supposed to be. On the one hand, we have seen that RQFT does predict that the appearances are consistent with the supposition that there are localized objects. So, for example, we could distinguish two record tokens at a given time if there were two disjoint regions O and O' and particle detector observables $C \in \mathcal{R}_\delta(O)$ and $C' \in \mathcal{R}_\delta(O')$ (approximated by observables *strictly* localized in O and O respectively) such that $\langle\psi, C\psi\rangle \approx 1$ and $\langle\psi, C'\psi\rangle \approx 1$. Now, it may be that Barrett is also worried about how, given a field ontology, we could assign any sort of trans-temporal identity to our record tokens. But this problem, however important philosophically, is distinct from the problem of localization. Indeed, the problem of trans-temporal identification of particles also arises in the context of non-relativistic QFT, where there is no problem of localization. Finally, Barrett might object that once we supply a quantum-theoretical model of a particle detector itself, then the superposition principle will prevent the field and detector from getting into a state where there is a fact of the matter as to whether a particle has been detected in the region O. But this is simply a restatement of the standard quantum measurement problem that infects *all* quantum theories — and we have made no pretense of solving that here.

10.8 Conclusion

Malament claims that his theorem justifies the belief that,

> ...in the attempt to reconcile quantum mechanics with relativity theory...one is driven to a field theory; all talk about "particles" has to be understood, at least in principle, as talk about the properties of, and interactions among,

quantized fields. (Malament 1996, 1)

In order to buttress Malament's argument for this claim, we provided two further results (Theorems 3 and 5) which show that the conclusion continues to hold for generic spacetimes, as well as for unsharp localization observables. We then went on to show that RQFT does not permit an ontology of localizable particles; and so, strictly speaking, our talk about localizable particles is a fiction. Nonetheless, RQFT does permit *talk* about particles — albeit, if we understand this talk as really being about the properties of, and interactions among, quantized fields. Indeed, modulo the standard quantum measurement problem, RQFT has no trouble explaining the appearance of macroscopically well-localized objects, and shows that our talk of particles, though a *façon de parler*, has a legitimate role to play in empirically testing the theory.

10.9 Appendix

Theorem 2 ***(Hegerfeldt).*** *Suppose that the localization system $(\mathcal{H}, \Delta \mapsto E_\Delta, t \mapsto U_t)$ satisfies monotonicity, time-translation covariance, energy bounded below, and NIWS. Then $U_t E_\Delta U_{-t} = E_\Delta$ for all $\Delta \subset \Sigma$ and all $t \in \mathbb{R}$.*

Proof. The formal proof corresponds directly to Hegerfeldt's informal proof. Thus, let Δ be a subset of some spatial hypersurface Σ. If $E_\Delta = 0$ then obviously $U_t E_\Delta U_{-t} = E_\Delta$ for all $t \in \mathbb{R}$. So, suppose that $E_\Delta \neq 0$, and let ψ be a unit vector such that $E_\Delta \psi = \psi$. Since Σ is a manifold, and since $\Delta \neq \Sigma$, there is a family $\{\Delta_n : n \in \mathbb{N}\}$ of subsets of Σ such that, for each $n \in \mathbb{N}$, the distance between the boundaries of Δ_n and Δ is nonzero, and such that $\bigcap_n \Delta_n = \Delta$. Fix $n \in \mathbb{N}$. By NIWS and time-translation covariance, there is an $\epsilon_n > 0$ such that $E_{\Delta_n} U_t \psi = U_t \psi$ whenever $0 \leq t < \epsilon_n$. That is, $\langle U_t \psi, E_{\Delta_n} U_t \psi \rangle = 1$ whenever $0 \leq t < \epsilon_n$. Since energy is bounded from below, we may apply Lemma 10.1 with $A = I - E_{\Delta_n}$ to conclude that $\langle U_t \psi, E_{\Delta_n} U_t \psi \rangle = 1$ for all $t \in \mathbb{R}$. That is, $E_{\Delta_n} U_t \psi = U_t \psi$ for all $t \in \mathbb{R}$. Since this holds for all $n \in \mathbb{N}$, and since (by monotonicity) $E_\Delta = \bigwedge_n E_{\Delta_n}$, it follows that $E_\Delta U_t \psi = U_t \psi$ for all $t \in \mathbb{R}$. Thus, $U_t E_\Delta U_{-t} = E_\Delta$ for all $t \in \mathbb{R}$. □

Lemma 10.2. *Suppose that the localization system $(\mathcal{H}, \Delta \mapsto E_\Delta, \mathbf{a} \mapsto U(\mathbf{a}))$ satisfies localizability, time-translation covariance, and no absolute velocity. Let Δ be a bounded spatial set. If $U(\mathbf{a})E_\Delta U(\mathbf{a})^* = E_\Delta$ for all timelike translations $\mathbf{a}$ of M, then $E_\Delta = 0$.*

Proof. By no absolute velocity, there is a pair $(\mathbf{a}, \mathbf{b})$ of timelike translations such that $\Delta + (\mathbf{a} - \mathbf{b})$ is in Σ and is disjoint from Δ. By time-translation covariance, we have,

$$E_{\Delta+(\mathbf{a}-\mathbf{b})} = U(\mathbf{a})U(\mathbf{b})^* E_\Delta U(\mathbf{b})U(\mathbf{a})^* = E_\Delta. \tag{10.9}$$

Thus, localizability entails that E_Δ is orthogonal to itself, and so $E_\Delta = 0$. □

Lemma 10.3 *Let $\{\Delta_n : n = 0, 1, 2, \ldots\}$ be a covering of Σ, and let $E = \bigvee_{n=0}^{\infty} E_{\Delta_n}$. If probability conservation and time-translation covariance hold, then $U_t E U_{-t} = E$ for all $t \in \mathbb{R}$.*

Proof. Since $\{\Delta_n + t : n \in \mathbb{N}\}$ is a covering of $\Sigma + t$, probability conservation entails that $\bigvee_n E_{\Delta_n + t} = E$. Thus,

$$U_t E U_{-t} = U_t \left[\bigvee_{n=0}^{\infty} E_{\Delta_n} \right] U_{-t} = \bigvee_{n=0}^{\infty} \left[U_t E_{\Delta_n} U_{-t} \right] \tag{10.10}$$

$$= \bigvee_{n=0}^{\infty} E_{\Delta_n + t} = E, \tag{10.11}$$

where the third equality follows from time-translation covariance. □

In order to prove the next result, we will need to invoke the following lemma from Borchers (1967).

Lemma 10.4 ***(Borchers).*** *Let $U_t = e^{itH}$, where H is a self-adjoint operator with spectrum bounded from below. Let E and F be projection operators such that $EF = 0$. If there is an $\epsilon > 0$ such that*

$$[E, U_t F U_{-t}] = 0, \qquad 0 \le t < \epsilon,$$

then $E U_t F U_{-t} = 0$ for all $t \in \mathbb{R}$.

Lemma 10.5 *Let $U_t = e^{itH}$, where H is a self-adjoint operator with spectrum bounded from below. Let $\{E_n : n = 0, 1, 2, \ldots\}$ be a family of projection operators such that $E_0 E_n = 0$ for all $n \ge 1$, and let $E = \bigvee_{n=0}^{\infty} E_n$.*

If $U_tEU_{-t} = E$ for all $t \in \mathbb{R}$, and if for each $n \geq 1$ there is an $\epsilon_n > 0$ such that

$$[E_0, U_tE_nU_{-t}] = 0, \qquad 0 \leq t < \epsilon_n, \tag{10.12}$$

then $U_tE_0U_{-t} = E_0$ for all $t \in \mathbb{R}$.

Proof. If $E_0 = 0$ then the conclusion obviously holds. Suppose then that $E_0 \neq 0$, and let ψ be a unit vector in the range of E_0. Fix $n \geq 1$. Using (10.12) and Borchers' lemma, it follows that $E_0U_tE_nU_{-t} = 0$ for all $t \in \mathbb{R}$. Then,

$$\|E_nU_{-t}\psi\|^2 = \langle U_{-t}\psi, E_nU_{-t}\psi\rangle = \langle \psi, U_tE_nU_{-t}\psi\rangle \tag{10.13}$$

$$= \langle E_0\psi, U_tE_nU_{-t}\psi\rangle = \langle \psi, E_0U_tE_nU_{-t}\psi\rangle = 0, \tag{10.14}$$

for all $t \in \mathbb{R}$. Thus, $E_nU_{-t}\psi = 0$ for all $n \geq 1$, and consequently, $[\bigvee_{n\geq 1} E_n]U_{-t}\psi = 0$. Since $E_0 = E - [\bigvee_{n\geq 1} E_n]$, and since (by assumption) $EU_{-t} = U_{-t}E$, it follows that

$$E_0U_{-t}\psi = EU_{-t}\psi = U_{-t}E\psi = U_{-t}\psi, \tag{10.15}$$

for all $t \in \mathbb{R}$. □

Theorem 3. *Suppose that the localization system $(\mathcal{H}, \Delta \mapsto E_\Delta, t \mapsto U_t)$ satisfies localizability, probability conservation, time-translation covariance, energy bounded below, and microcausality. Then $U_tE_\Delta U_{-t} = E_\Delta$ for all Δ and all $t \in \mathbb{R}$.*

Proof. Let Δ be an open subset of Σ. If $\Delta = \Sigma$ then probability conservation and time-translation covariance entail that $E_\Delta = E_{\Delta+t} = U_tE_\Delta U_{-t}$ for all $t \in \mathbb{R}$. If $\Delta \neq \Sigma$ then, since Σ is a manifold, there is a covering $\{\Delta_n : n \in \mathbb{N}\}$ of $\Sigma\backslash\Delta$ such that the distance between Δ_n and Δ is nonzero for all n. Let $E_0 = E_\Delta$, and let $E_n = E_{\Delta_n}$ for $n \geq 1$. Then 1 entails that $E_0E_n = 0$ when $n \geq 1$. If we let $E = \bigvee_{n=0}^{\infty} E_n$ then probability conservation entails that $U_tEU_{-t} = E$ for all $t \in \mathbb{R}$ (see Lemma 10.3). By time-translation covariance and microcausality, for each $n \geq 1$ there is an $\epsilon_n > 0$ such that

$$[E_0, U_tE_nU_{-t}] = 0, \qquad 0 \leq t < \epsilon_n. \tag{10.16}$$

Since the energy is bounded from below, Lemma 10.5 entails that $U_tE_0U_{-t} = E_0$ for all $t \in \mathbb{R}$. That is, $U_tE_\Delta U_{-t} = E_\Delta$ for all $t \in \mathbb{R}$. □

Theorem 5. *Suppose that the unsharp localization system* $(\mathcal{H}, \Delta \mapsto A_\Delta, \mathbf{a} \mapsto U(\mathbf{a}))$ *satisfies additivity, translation covariance, energy bounded below, microcausality, and no absolute velocity. Then* $A_\Delta = 0$ *for all* Δ.

Proof. We prove by induction that $\|A_\Delta\| \leq (2/3)^m$, for each $m \in \mathbb{N}$, and for each bounded Δ. For this, let F_Δ denote the spectral measure for A_Δ.

(Base case: $m = 1$) Let $E_\Delta = F_\Delta(2/3, 1)$. We verify that $(\mathcal{H}, \Delta \mapsto E_\Delta, \mathbf{a} \mapsto U(\mathbf{a}))$ satisfies the conditions of Malament's theorem. Clearly, no absolute velocity and energy bounded below hold. Moreover, since unitary transformations preserve spectral decompositions, translation covariance holds; and since spectral projections of compatible operators are also compatible, microcausality holds. To see that localizability holds, let Δ and Δ' be disjoint bounded subsets of a single hyperplane. Then microcausality entails that $[A_\Delta, A_{\Delta'}] = 0$, and therefore $E_\Delta E_{\Delta'}$ is a projection operator. Suppose for reductio ad absurdum that ψ is a unit vector in the range of $E_\Delta E_{\Delta'}$. By additivity, $A_{\Delta\cup\Delta'} = A_\Delta + A_{\Delta'}$, and we therefore obtain the contradiction:

$$1 \geq \langle\psi, A_{\Delta\cup\Delta'}\psi\rangle = \langle\psi, A_\Delta\psi\rangle + \langle\psi, A_{\Delta'}\psi\rangle \geq 2/3 + 2/3 . \qquad (10.17)$$

Thus, $E_\Delta E_{\Delta'} = 0$, and Malament's theorem entails that $E_\Delta = 0$ for all Δ. Therefore, $A_\Delta = A_\Delta F_\Delta(0, 2/3)$ has spectrum lying in $[0, 2/3]$, and $\|A_\Delta\| \leq 2/3$ for all bounded Δ.

(Inductive step) Suppose that $\|A_\Delta\| \leq (2/3)^{m-1}$ for all bounded Δ. Let $E_\Delta = F_\Delta((2/3)^m, (2/3)^{m-1})$. In order to see that Malament's theorem applies to $(\mathcal{H}, \Delta \mapsto E_\Delta, \mathbf{a} \mapsto U(\mathbf{a}))$, we need only check that localizability holds. For this, suppose that Δ and Δ' are disjoint subsets of a single hyperplane. By microcausality, $[A_\Delta, A_{\Delta'}] = 0$, and therefore $E_\Delta E_{\Delta'}$ is a projection operator. Suppose for reductio ad absurdum that ψ is a unit vector in the range of $E_\Delta E_{\Delta'}$. Since $\Delta \cup \Delta'$ is bounded, the induction hypothesis entails that $\|A_{\Delta\cup\Delta'}\| \leq (2/3)^{m-1}$. By additivity, $A_{\Delta\cup\Delta'} = A_\Delta + A_{\Delta'}$, and therefore we obtain the contradiction:

$$(2/3)^{m-1} \geq \langle\psi, A_{\Delta\cup\Delta'}\psi\rangle = \langle\psi, A_\Delta\psi\rangle + \langle\psi, A_{\Delta'}\psi\rangle \geq (2/3)^m + (2/3)^m . \qquad (10.18)$$

Thus, $E_\Delta E_{\Delta'} = 0$, and Malament's theorem entails that $E_\Delta = 0$ for all Δ. Therefore, $\|A_\Delta\| \leq (2/3)^m$ for all bounded Δ. □

Theorem 6. *Suppose that the system* $(\mathcal{H}, \Delta \mapsto N_\Delta, \mathbf{a} \mapsto U(\mathbf{a}))$ *of local number operators satisfies additivity, translation covariance, energy bounded* *below, number conservation, microcausality, and no absolute velocity. Then, $N_\Delta = 0$ for all bounded Δ.*

Proof. Let N be the unique total number operator obtained from taking the sum $\sum_n N_{\Delta_n}$ where $\{\Delta_n : n \in \mathbb{N}\}$ is a disjoint covering of Σ. Note that for any $\Delta \subseteq \Sigma$, we can choose a covering containing Δ, and hence, $N = N_\Delta + A$, where A is a positive operator. By microcausality, $[N_\Delta, A] = 0$, and therefore $[N_\Delta, N] = [N_\Delta, N_\Delta + A] = 0$. Furthermore, for any vector ψ in the domain of N, $\langle \psi, N_\Delta \psi \rangle \leq \langle \psi, N\psi \rangle$.

Let E be the spectral measure for N, and let $E_n = E(0, n)$. Then, NE_n is a bounded operator with norm at most n. Since $[E_n, N_\Delta] = 0$, it follows that

$$\langle \psi, N_\Delta E_n \psi \rangle = \langle E_n \psi, N_\Delta E_n \psi \rangle \leq \langle E_n \psi, N E_n \psi \rangle \leq n \,, \tag{10.19}$$

for any unit vector ψ. Thus, $\|N_\Delta E_n\| \leq n$. Since $\bigcup_{n=1}^{\infty} E_n(\mathcal{H})$ is dense in $\mathcal{H}$, and since $E_n(\mathcal{H})$ is contained in the domain of N_Δ (for all n), it follows that if $N_\Delta E_n = 0$, for all n, then $N_\Delta = 0$. We now concentrate on proving the antecedent.

For each Δ, let $A_\Delta = (1/n) N_\Delta E_n$. We show that the structure $(\mathcal{H}, \Delta \mapsto A_\Delta, \mathbf{a} \mapsto U(\mathbf{a}))$ satisfies the conditions of Theorem 5. Clearly, energy bounded below and no absolute velocity hold. It is also straightforward to verify that additivity and microcausality hold. To check translation covariance, we compute:

$$U(\mathbf{a}) A_\Delta U(\mathbf{a})^* = U(\mathbf{a}) N_\Delta E_n U(\mathbf{a})^* = U(\mathbf{a}) N_\Delta U(\mathbf{a})^* U(\mathbf{a}) E_n U(\mathbf{a})^* \tag{10.20}$$

$$= U(\mathbf{a}) N_\Delta U(\mathbf{a})^* E_n = N_{\Delta + \mathbf{a}} E_n = A_{\Delta + \mathbf{a}}. \tag{10.21}$$

The third equality follows from number conservation, and the fourth equality follows from translation covariance. Thus, $N_\Delta E_n = A_\Delta = 0$ for all Δ. Since this holds for all $n \in \mathbb{N}$, $N_\Delta = 0$ for all Δ. □

Acknowledgements

Thanks to Jeff Barrett and David Malament for helpful comments on an earlier version of this paper.

References

Barrett, Jeffrey A. (2001): On the nature of measurement records in relativistic quantum field theory. This volume.

Borchers, H.-J. (1967): A remark on a theorem of B. Misra. *Communications in Mathematical Physics* 4: 315–323.

Buchholz, Detlev (2000): Algebraic quantum field theory: A status report. E-Print: arXiv.org math-ph/0011044.

Busch, Paul (1999): Unsharp localization and causality in relativistic quantum theory. *Journal of Physics A* 32: 6535–6546.

Chaiken, Jan M. (1968): Number operators for representations of the canonical commutation relations. *Communications in Mathematical Physics* 8: 164–184.

Cushing, James T. (1996): What measurement problem? In R. Clifton (ed.), *Perspectives on Quantum Reality.* Dordrecht: Kluwer, 167–181.

Dickson, W. Michael (1998): *Quantum Chance and Nonlocality.* Cambridge: Cambridge University Press.

Earman, John (1986): *A Primer on Determinism.* Boston: Reidel.

Fleming, Gordon, and Jeremy Butterfield (1999): Strange positions. In J. Butterfield and C. Pagonis (eds.), *From Physics to Philosophy.* Cambridge: Cambridge University Press, 108–165.

Haag, Rudolf (1992): *Local Quantum Physics.* Berlin: Springer.

Halvorson, Hans (2001): Reeh-Schlieder defeats Newton-Wigner: On alternative localization schemes in relativistic quantum field theory, *Philosophy of Science* 68: 111–133.

Hegerfeldt, Gerhard C. (1998a): Causality, particle localization and positivity of the energy. In A. Böhm, et al. (eds.), *Irreversibility and Causality.* Berlin: Springer, 238–245.

Hegerfeldt, Gerhard C. (1998b): Instantaneous spreading and Einstein causality in quantum theory. *Annalen der Physik* 7: 716–725.

Holland, Peter R. (1993): *The Quantum Theory of Motion.* Cambridge: Cambridge University Press.

Malament, David (1996): In defense of dogma: Why there cannot be a relativistic quantum mechanics of (localizable) particles. In R. Clifton (ed.), *Perspectives on Quantum Reality.* Dordrecht: Kluwer, 1–10.

Maudlin, Tim (1994): *Quantum Non-Locality and Relativity.* Cambridge: Blackwell.

Redhead, Michael (1995): The vacuum in relativistic quantum field theory. In D. Hull, M. Forbes, and R. M. Burian (eds.), *PSA 1994*, Vol. 2. East Lansing, MI: Philosophy of Science Association, 77–87.

Requardt, Manfred (1982): Spectrum condition, analyticity, Reeh-Schlieder and cluster properties in non-relativistic Galilei-invariant quantum theory. *Journal of Physics A* 15: 3715–3723.

Schlieder, S. (1971): Zum kausalen Verhalten eines relativistischen quantenmechanischen Systems. In H.-.P. Dürr (ed.), *Quanten und Felder.* Braunschweig: Vieweg, 145–160.

Segal, Irving E. (1964): Quantum fields and analysis in the solution manifolds of differential equations. In W. T. Martin and I. E. Segal, (eds.), *Proceedings of a Conference on the Theory and Applications of Analysis in Function Space.* Cambridge: MIT Press, 129–153.

Streater, Raymond F. and Arthur S. Wightman (2000): *PCT, Spin and Statistics, and All That* (3rd edition). Princeton: Princeton University Press.

Thaller, Bernd (1992): *The Dirac Equation.* Berlin: Springer.

Ticiatti, Robin (1999): *Quantum Field Theory for Mathematicians.* Cambridge: Cambridge University Press.

This page is intentionally left blank

Chapter 11

Events and Covariance in the Interpretation of Quantum Field Theory

Dennis Dieks
Utrecht University

Abstract. *In relativistic quantum field theory the notion of a local* operation *is basic: each open space-time region is associated with an algebra of observables representing possible measurements performed within this region. It is much more difficult to accommodate the notions of* events *taking place in such regions or of localized objects. But how could the notion of a local operation be basic in the theory if this same theory were not able to represent localized measuring devices and localized events? After briefly reviewing these difficulties we discuss a strategy for eliminating the tension, namely by interpreting quantum theory in a realist way. To implement this strategy we use the ideas of the modal interpretation of quantum mechanics. We then consider the question of whether the resulting scheme can be made Lorentz-invariant.*

11.1 The Problem

Relativistic quantum theory is notorious for the difficulties it has with the concept 'position'. Relativistic quantum mechanics does not accommodate position as a particle observable in a natural way; the Newton-Wigner position operator, which is the only observable that comes into serious consideration, faces well-known difficulties in connection with covariance. The standard response is that this signals the inadequacy of the particle concept in relativistic quantum theory, and that we should switch to quantum *field* theory. The most general form of this theory is algebraic relativistic quantum field theory (ARQFT). At first sight, the prospects for position as a central physical magnitude seem very good in this theory. Indeed, ARQFT is formulated against the background of Minkowski space-time, and regions of Minkowski space-time figure prominently in the axioms of the theory. Specifically, a $C^{\star}$-algebra of observables is associated with each open region O of Minkowski space-time. These operators represent physical operations, measurements, that can be performed within O (Haag 1992, p. 105). The localization of observables is thus a fundamental notion in the theory. However, the localization of an event or an object are troublesome notions in ARQFT. We already mentioned that the notion of a localized system is problematic in relativistic quantum mechanics (Hegerfeldt 1974, 1985, Malament 1996), and it turns out that similar problems persist in field theory.

Within the field theoretical framework the notorious Reeh-Schlieder theorem is responsible for additional concern about locality. According to this theorem (Haag 1992, p. 101), the set of states obtained by applying the local operations associated with any particular region O to a state of bounded energy is dense in the Hilbert space of all states of the total field. That means that any state can be approximated, arbitrarily closely, by applying local operations to the vacuum (or any other state of bounded energy). So, even states that could be thought of as candidates for representing objects far from O can be generated by local operations within O. A corollary is that the local $C^{\star}$-algebras do not contain observables corresponding to tests of whether or not the vacuum state is the state of the total field (Haag 1992, p. 102). In other words, even in the vacuum state local tests (corresponding to local observables) of whether or not matter, energy or charge are present will with some probability yield positive results—particle counters measuring these local quantities will click from time to time. The vacuum can therefore not be thought of as a 'sum of local vacua'; it is an inherently global concept. The same applies to N-particle states: it is impossible to verify the presence of such a state by local operations. The particles of the N-particle states are not local objects.

The Reeh-Schlieder theorem demonstrates that the vacuum, and all other states of bounded energy, have long-distance correlations built into them. It is therefore not surprising to find that Bell inequalities are violated in these states—a standard sign of non-locality.

However, one should be careful not to lump all such non-local features together. That there are non-local correlations does not mean that it is impossible to speak about localized objects; on the contrary, it only makes sense to discuss non-local correlations if there are more or less localized things, at large mutual distances, that can be correlated with each other. And a natural conclusion to draw from the Hegerfeldt and Malament results is that relativistic quantum theory does not admit states that assign non-zero probabilities to a bounded region only. But that is not in *a priori* conflict with the idea that there are localized events or localized objects; there could be very many of them, (almost) everywhere.

In this paper we will concentrate on the question of whether it is possible to interpret algebraic quantum field theory in such a way that the theory is able to accommodate the concept of an event localized in a small spacetime region, and whether it is possible to work—at least in some circumstances—with the concept of a localized physical system. We will not address the ramifications of violations of the Bell-inequalities and similar features of non-locality relating to long-distance correlations.

The question of how to handle localized systems is vital for the interpretation of relativistic quantum field theory. Without an answer to it the theory is incomplete at best and inconsistent at worst. For how can the local operations whose existence the theory assumes be performed if the theory itself does not allow the description of localized devices? If the treatment of such devices is outside the scope of the theory, the theory is incomplete in its description of physical reality. Even more worrying would be the alternative that the theory is complete and still cannot handle localized systems. That would signal inconsistency, given the fundamental importance of localized operations in the theory.

A closely related issue concerns the meaning of the Minkowski background that is assumed in the formulation of ARQFT. Of course, we think of this four-dimensional manifold as the space-time arena in which physical processes take place. However, without the existence of at least approximately localized objects or events the usual operational meaning and material realization of space-time regions would no longer be available. It is not clear that the four-dimensional manifold would in this case still have its usual significance as space-time. Notions like causality between events and the spacetime distance between events, which figure in the axioms of ARQFT, would not have the physical correlates they are commonly assumed to possess (e.g., in terms of signals connecting the events). That would make the physical motivation of these axioms (like the micro-causality axiom) uncertain.

11.2 What Has to Be Done

What we would like to have, in order to achieve a consistent and possibly complete theory, is a representation of the physical systems performing the local operations mentioned in the axioms. This would make it possible to consider the concepts of measurement and operation not as fundamental but as derived: a local measurement would be an interaction between an object and a localized measuring device. The natural way to achieve this is to describe measuring devices in terms of 'beables', that is as systems characterized by the values certain physical magnitudes assume in them. Clearly, a prerequisite for the successful implementation of this program is that we have an interpretation of the formalism of ARQFT in terms of physical magnitudes and the values that they can take; as opposed to the usual interpretation, which is only about measurement results. This is nothing else than a version of the general problem of the interpretation of quantum theory.

Although the traditional interpretational problems are rarely discussed in the context of quantum field theory, it is clear that they exist there in analogous form and are no less urgent than in non-relativistic quantum mechanics. As in the non-relativistic theory, the central issue is that it is not obvious that the theory is about objective physical states of affairs, even in circumstances in which no macroscopic measurements are being made. This is because the fields in quantum field theory do not attach values of physical magnitudes to spacetime points. Rather, they are fields of operators, with a standard interpretation in terms of macroscopic measurement results. As just pointed out, we would like to give another meaning to the formalism, namely in terms of physical systems that possess certain properties. What we would like to do is to provide an interpretation in which not only operators, but also *properties* are assigned to spacetime regions. That is, we would like at least some of the observables to take on definite values. This would lead to a picture in which it is possible to speak of objective *events* (if some physical magnitude takes on a definite value in a certain spacetime region, this constitutes the event that this magnitude has that value there and then).

As a second step, we need to show that it is possible to have consistent joint probabilities of the properties assigned to different spacetime regions. Finally, it should be made clear that it is possible, at least in principle, to assign properties to spacetime regions in such a way that (approximately) localized systems can be represented. This does not mean that it is necessary to establish that there exist localized systems in every physical state allowed by the theory. The Minkowski manifold need not have the space-time implementation that we are used to in every possible world allowed by the theory. Rather, we have to show that localized systems exist in some states; and in particular, we expect localized systems to emerge in a classical limiting situation.

Obviously, the above programme is ambitious and addresses complicated and controversial issues. On several points we will not be able to do much more than suggesting possibilities and indicating open questions.

11.3 Modal Property Attribution Schemes

In order to be able to speak about events and physical properties of systems rather than merely about the results of operations and measurements, we will make an attempt to apply modal interpretational ideas to ARQFT. Modal interpretations of quantum mechanics (van Fraassen 1982, Dieks 1989, Healey 1989, Dieks and Vermaas 1998) interpret the mathematical formalism of quantum theory in terms of properties possessed by physical systems, i.e. quantum mechanical observables taking on definite values. Because of the Kochen and Specker no-go theorem, not all observables pertaining to a system can be definite-valued at the same time. Modal interpretations therefore specify a *subset* of all observables, such that only the observables in this subset are definite-valued. It is characteristic of the modal approach in the version that we shall use that this is done in a state-dependent way: the quantum mechanical state of the system contains all information needed to determine the set of definite-valued observables.

In non-relativistic quantum mechanics, and in the most common version of the modal interpretation (Vermaas and Dieks 1995), the precise prescription for finding this set makes use of the spectral decomposition of the density operator of the system, in the following way. Let α be our system and let β represent its total environment (the rest of the universe). Let $\alpha\&\beta$ be represented by $|\psi^{\alpha\beta}\rangle \in \mathcal{H}^\alpha \otimes \mathcal{H}^\beta$. The bi-orthonormal decomposition of $|\psi^{\alpha\beta}\rangle$,

$$|\psi^{\alpha\beta}\rangle = \sum_i c_i \, |\psi_i^\alpha\rangle \, |\psi_i^\beta\rangle, \tag{11.1}$$

with $\langle\psi_i^\alpha|\psi_j^\alpha\rangle = \langle\psi_i^\beta|\psi_j^\beta\rangle = \delta_{ij}$, generates a set of projectors operating on $\mathcal{H}^\alpha$: $\{|\psi_i^\alpha\rangle\langle\psi_i^\alpha|\}_i$. If there is no degeneracy among the numbers $\{|c_i|^2\}$, this is a uniquely determined set of one-dimensional projectors. If there is degeneracy, the projectors belonging to one value of $\{|c_i|^2\}$ can be added to form a multi-dimensional projector; the thus generated new set of projectors, including multi-dimensional ones, is again uniquely determined. These projectors are the ones occurring in the spectral decomposition of the reduced density operator of α.

The modal interpretation of non-relativistic quantum mechanics assigns definite values to the subset of all physical magnitudes that are generated by these projectors; i.e., the subset obtained by starting with these projectors, and then including their continuous functions, real linear combinations, symmetric and antisymmetric products, and finally closing the set (Clifton 1996) (the thus defined subset of all observables constitutes the set of 'well-defined' or 'applicable' physical magnitudes, in Bohrian parlance). *Which* value among the possible values of a definite magnitude is actually realized is not fixed by the interpretation. For each possible value a probability is specified: the probability that the magnitude represented by $|\psi_i^\alpha\rangle\langle\psi_i^\alpha|$ has the value 1 is given by $|c_i|^2$. In the case of degeneracy it is stipulated that the magnitude represented by $\sum_{i\in I_l}|\psi_i^\alpha\rangle\langle\psi_i^\alpha|$ has value 1 with probability $\sum_{i\in I_l}|c_i|^2$ (I_l is the index-set containing indices j, k such that $|c_j|^2 = |c_k|^2$).

The observation that the definite-valued projections occur in the spectral decomposition of α's density operator gives rise to a generalization of the above scheme that is also applicable to the case in which the total system $\alpha\&\beta$ is not represented by a pure state: find α's density operator by partial tracing from the total density operator, determine its spectral resolution and construct the set of definite-valued observables from the projection operators in this spectral resolution (Vermaas and Dieks 1995).

The above recipe for assigning properties is meant to apply to each physical system in a non-overlapping collection of systems that together make up the total universe (Bacciagaluppi and Dickson 1999, Dieks 1998a). It is easy to write down a satisfactory *joint* probability distribution for the properties of such a collection (or a subset of it):

$$Prob(P_i^\alpha, P_j^\beta, ..., P_k^\theta, ..., P_l^\xi) = \langle\Psi|P_i^\alpha.P_j^\beta.....P_k^\theta.....P_l^\xi|\Psi\rangle, \tag{11.2}$$

where the left-hand side represents the joint probability for the projectors occurring in the argument of taking the value 1, and where Ψ is the state of the total system consisting of α, β, θ, etc. (Vermaas and Dieks 1995). It is important for the consistency of this probability ascription that the projection operators occurring in the formula all commute (which they do, since they operate in different Hilbert spaces).

11.4 A Perspectival Version of the Modal Interpretation

In the usual version of the modal interpretation, as it was just explained, physical properties are represented by *monadic* predicates. Such properties belong to a system without reference to other systems. We will now briefly describe an alternative, according to which physical properties have a *relational* character (Bene and Dieks 2001). These relational properties need *two* systems for their definition: the physical system S and a 'reference system' R that defines the 'perspective' from which S is considered. This reference system is a larger system, of which S is a part. We will allow that one and the same system, at one and the same instant of time, can have different states with respect to different reference systems. However, the system will have one single state with respect to any given reference system. This state of S with respect to R is a density matrix denoted by ρ_R^S. In the special case in which R coincides with S, we have the 'state of S with respect to itself', which we take to be the same as the state of S assigned by the modal scheme explained in the previous section; i.e. it is one of the projectors occurring in the spectral decomposition of the reduced density operator of S.

The rules for determining all states, for arbitrary S and R, are as follows. If U is the whole universe, then ρ_U^U is taken as the quantum state assigned to U by standard quantum theory. If system S is contained in system A, the state ρ_A^S is defined as the density operator that can be derived from ρ_A^A by taking the partial trace over the degrees of freedom in A that do not pertain to S:

$$\rho_A^S = \mathrm{Tr}_{A\setminus S}\, \rho_A^A \tag{11.3}$$

Any relational state of a system with respect to a bigger system containing it can be derived by means of Eq. (11.3).

As in the ordinary modal scheme, the state ρ_U^U evolves unitarily in time. Because there is no collapse of the wave function in the modal interpretation, this unitary evolution of the total quantum state is the main dynamical principle of the theory. Furthermore, it is assumed that the state assigned to a *closed* system S undergoes a unitary time evolution

$$i\hbar\frac{\partial}{\partial t}\rho_S^S = [H_S,\, \rho_S^S] \tag{11.4}$$

As always in the modal interpretation, the idea is that the theory should specify only the *probabilities* of the various possibilities. For a collection of pair-wise disjoint systems, with respect to one reference system, one could postulate that the joint probability of the states of the various systems is given by the usual formula, Eq. (11.2) (Bene and Dieks 2001 uses a different probability postulate). A significant point is that joint probabilities should not always be expected to exist within the perspectival approach because states that are defined with respect to different quantum reference systems need not be commensurable.

The Schmidt representation of the total state shows that the states of A and its complement $U \setminus A$, with respect to themselves, are one-to-one correlated. Therefore, knowledge of the state of $U \setminus A$, plus the total state, makes it possible to infer the state of A. This suggests that one may consider the state of S with respect to the reference system A, ρ_A^S, alternatively as being defined *from the perspective* $U \setminus A$ (here A is an arbitrary quantum reference system, while U is the whole universe). Sometimes the concept of a 'perspective' is intuitively more appealing than the concept of a quantum reference system (cf. Rovelli 1996). However, the notion of a perspective has some limitations. First, if A itself is the whole universe, the concept of an external perspective cannot be applied. Moreover, the state of the system $U \setminus A$ in itself does not contain sufficient information to determine the state of system A; one also needs the additional information provided by $|\psi_U>$ in order to compute $|\psi_A>$. But $|\psi_A>$ does contain all the information needed to calculate ρ_A^S (cf. Eq. (11.3)). We will therefore relativize the states of S to reference systems that contain S, although we shall sometimes—in cases in which this is equivalent— also speak about the state of S *from the perspective* of the complement of the reference system.

Of course, we must address the question of the physical meaning of the states ρ_A^S. In the perspectival approach it is a fundamental assumption that basic descriptions of the physical world have a relational character, and therefore we cannot explain the relational states by appealing to a definition in terms of more basic, and more familiar, non-relational states. But we should at the very least explain how these relational states connect to actual experience. Minimally, the theory has to give an account of what observers observe. We postulate that experience in this sense is represented by the state of a part of the observer's perceptual apparatus (the part characterized by a relevant indicator variable, like the display of a measuring device) *with respect to itself*. More generally, the states of systems with respect to themselves correspond to the (monadic)properties assigned by the earlier, non-perspectival, version of the modal interpretation. The empirical meaning of many other states can be understood and explained—by using the rules of the interpretation—through their relation to these states of observers, measuring devices, and other systems, with respect to themselves. For more details concerning these ideas see Bene and Dieks 2001.

11.5 Application to ARQFT

The modal interpretations that were discussed in the previous sections were devised for the case of quantum mechanics, in which each physical system is represented within its own Hilbert space and in which the total Hilbert space is the tensor product of the Hilbert spaces of the individual component systems. The possible physical properties of the systems correspond to observables defined as operators on the Hilbert spaces of these systems. In axiomatic quantum field theory algebras of observables are associated with open spacetime regions ('local algebras'). It is therefore natural to think of these observables as representative of possible local physical characteristics and to regard the spacetime regions as the analogues of the physical systems to which we applied the modal scheme before. If the open spacetime regions would correspond to subspaces of the total Hilbert space, an immediate application of the modal scheme would be possible and would lead to the selection of definite-valued observables from the local algebras. Unfortunately, things are not that simple. The local algebras must generally be expected to be of type III (algebra's with only infinite-dimensional projectors, without minimal projectors). This implies that they cannot be represented as algebras of bounded observables on a Hilbert space (such algebras are of type I). In other words, the local algebras cannot be thought of as algebras of observables belonging to a physical system described within its own Hilbert space, and the total Hilbert space is not a tensor product of Hilbert spaces of such local subsystems.

There are, however, other ways to introduce the notion of a localized subsystem. One possibility is to use the algebras of type I that 'lie between two local algebras'. That such type-I algebras exist is assumed in the postulate of the 'split property' (Haag 1992, Ch. V.5). If this postulate is accepted one can consider the algebras of type I lying between the $C^\star$-algebras associated with concentric standard 'diamond' regions with radii r and $r + \epsilon$, respectively, with r and ϵ very small numbers. In this way we approximate the notion of a spacetime point as a physical system, represented in a subspace of the total Hilbert space (Dieks 2000). The advantage of this approach is that all systems that we consider have their own Hilbert spaces, and that we can therefore use the same techniques as in non-relativistic quantum mechanics. However, a disadvantage is the arbitrariness in fixing the values of r and ϵ, and in choosing one type-I algebra from the infinity of such algebras lying between the two type-III algebras associated with the two diamond regions (Clifton 2000).

Another way of applying the modal ideas to ARQFT was suggested by Clifton (2000); in this proposal there is much less arbitrariness. Clifton considers an arbitrary von Neumann algebra $\mathcal{R}$. A state ρ on $\mathcal{R}$ defines the 'centralizer subalgebra'

$$\mathcal{C}_{\rho,\mathcal{R}} \equiv \{A \in \mathcal{R} : \rho([A, B]) = 0 \text{ for all } B \in \mathcal{R}\}. \tag{11.5}$$

Further, let $\mathcal{Z}(\mathcal{C}_{\rho,\mathcal{R}})$ be the *center algebra* of $\mathcal{C}_{\rho,\mathcal{R}}$, i.e. the elements of $\mathcal{C}_{\rho,\mathcal{R}}$ that commute with all elements of $\mathcal{C}_{\rho,\mathcal{R}}$. Clifton proves the following theorem:

Let $\mathcal{R}$ be a von Neumann algebra and ρ a faithful normal state of $\mathcal{R}$ with centralizer $\mathcal{C}_{\rho,\mathcal{R}} \subseteq \mathcal{R}$. Then $\mathcal{Z}(\mathcal{C}_{\rho,\mathcal{R}})$, the center of $\mathcal{C}_{\rho,\mathcal{R}}$, is the unique subalgebra $\mathcal{S} \subseteq \mathcal{R}$ such that:

(1) *The restriction of ρ to $\mathcal{S}$ is a mixture of dispersion-free states.*
(2) *$\mathcal{S}$ is definable solely in terms of ρ and the algebraic structure of $\mathcal{R}$.*
(3) *$\mathcal{S}$ is maximal with respect to both just-mentioned properties.*

Moreover, for faithful states projections from $\mathcal{Z}(\mathcal{C}_{\rho,\mathcal{R}})$ are strictly correlated with projections from $\mathcal{Z}(\mathcal{C}_{\rho,\mathcal{R}'})$, where $\mathcal{R}'$ is the commutant of $\mathcal{R}$. This generalizes the strict correlation between projectors occurring in the bi-orthonormal decomposition in quantum mechanics (Clifton 2000).

These results make it natural to take $\mathcal{Z}(\mathcal{C}_{\rho,\mathcal{R}(\Diamond_r)})$ as the subalgebra of definite-valued observables in $\mathcal{R}(\Diamond_r)$, the algebra associated with a diamond region $\Diamond_r$, if ρ is a pure state of the field that induces a faithful state on $\mathcal{R}(\Diamond_r)$ (for example, ρ could be the vacuum or any other state with bounded energy). In this way it becomes possible to assign definite physical properties to regions of spacetime.

As in the case of non-relativistic quantum mechanics, we will take the projectors in the (Abelian) algebra of definite-valued observables as a 'base set' of definite-valued quantities. The complete collection of definite-valued observables can be constructed from this base set by closing the set under the operations of taking continuous functions, real linear combinations, and symmetric and anti-symmetric products (Clifton 1998). The probability of projector P_l having the value 1 is $\langle\Psi|P_l|\Psi\rangle$, with $|\Psi>$ the state of the total field. Subdividing Minkowski spacetime into a collection of non-overlapping point-like regions, and applying the above prescription to the associated algebras, we achieve the picture aimed at: to each spacetime region belong definite values of some physical magnitude, and this constitutes an event localized in that region.

It is important, though, to realize that the resulting picture is not classical. A typical quantum feature is that there is no guarantee that the definite-valued quantities associated with a spacetime region will also be definite-valued quantities of larger spacetime regions that contain the original one. This is analogous to what is found in non-relativistic quantum mechanics: in general, the properties of systems do not follow from the properties of their components. That such a simple relation between wholes and parts nevertheless does obtain in classically describable situations needs to be explained by some physical mechanism; decoherence is the prime candidate. Indeed, it is easy to see that the principle of 'property composition' (asserting that the properties of composite systems can be built up from the properties of their components) holds according to the modal interpretation of non-relativistic quantum mechanics if the environment 'decoheres' all component systems separately. For a composite system $\alpha\&\beta$, with environments E^α and E^β of α and β, respectively, we have in this case the state

$$|\psi\rangle = \sum_{ij} c_{ij}\, |\psi_i^\alpha\rangle\, |E_i^\alpha\rangle\, |\psi_j^\beta\rangle\, |E_j^\beta\rangle, \tag{11.6}$$

with $\langle\psi_i^\alpha|\psi_j^\alpha\rangle = \langle\psi_i^\beta|\psi_j^\beta\rangle = \langle E_i^\alpha|E_j^\alpha\rangle = \langle E_i^\beta|E_j^\beta\rangle = \delta_{ij}$. The definite-valued projectors for $\alpha\&\beta$ are therefore the products of the definite-valued projectors for α and β separately.

In the field-theoretic case decoherence similarly tends to mask typical quantum effects. Above, we discussed two possible ways of implementing modal ideas in the field-theoretic context. If the split property is used, physical subsystems are represented by Hilbert spaces, and the same reasoning can be employed as in the non-relativistic case. If the algebras of definite-valued observables as defined by Clifton are taken as a starting point, the situation does not become very different. These algebras are generated by projection operators P that are strictly correlated to projections $\overline{P}$ associated with the environment: $< \Psi|P\overline{P}|\Psi > = < \Psi|P|\Psi > = < \Psi|\overline{P}|\Psi >$. Now suppose that decoherence mechanisms have been effective, in the sense that information has been carried away to distant regions, and that (semi-)permanent memories have been formed of the definite-valued observables in the spacetime region O. If there are thus 'copy' projectors $\overline{P}$, strictly correlated to P, that belong to far-away regions, we have that $\rho([P,B]) = \rho([\overline{P},B] = 0$, for all observables B associated with a region O' that contains O but is not too big (so that $\overline{P}$ belongs to an algebra that is distant enough to be in the commutant of the observables B). Therefore, the centralizer subalgebra of the von Neumann algebra associated with O will be contained in the centralizer subalgebra belonging to O'. It follows that the definite-valued projectors associated with region O commute with those associated with O' (the latter by definition commute with all elements of the centralizer subalgebra of O', and therefore with all elements of the centralizer subalgebra of O). The properties of O and O' are therefore compatible.

11.6 Histories of Modal Properties

Let us return to the non-relativistic case to consider the problem of correlations in time. There is a natural analogue of expression (11.2) for the case of Heisenberg projection operators pertaining to different instants of time:

$$Prob(P_i(t_1), P_j(t_2), ...P_l(t_n)) = \langle\Psi|P_i(t_1).P_j(t_2).....P_l(t_n).P_l(t_n).....P_j(t_2).P_i(t_1)|\Psi\rangle. \quad (11.7)$$

This expression is in accordance with the standard prescription for calculating the joint probability of outcomes of consecutive measurements. It agrees also with the joint distribution assigned to 'consistent histories' in the consistent histories approach to the interpretation of quantum mechanics (Griffiths 1984). However, it should be noted that the projection operators in (11.7), pertaining to different times as they do, need not commute. As a result, (11.7) does not automatically yield a consistent probability distribution. For this reason it is an essential part of the consistent histories approach to impose the following 'decoherence condition', in order to guarantee that expression (11.7) is an ordinary Kolmogorov probability:

$$\begin{aligned} \langle\Psi|P_i(t_1).P_j(t_2).....P_l(t_n).P_{l'}(t_n)......P_{j'}(t_2).P_{i'}(t_1)|\Psi\rangle &= 0, \\ \text{if} i \neq i' \vee j \neq j' \vee .... \vee l \neq l'. & \end{aligned} \tag{11.8}$$

In the consistent histories approach the only sequences of properties which are considered are those satisfying the decoherence condition (11.8). It has been argued in the literature (Kent 1995) that the projection operators singled out by the modal interpretation will in general not satisfy this decoherence condition. That argument is not valid, however (Dieks 2000). On the contrary, it is natural to introduce the idea of decoherence in the modal scheme in such a way that condition (11.8) is satisfied. Eq. (11.7) yields a consistent joint multi-times probability distribution for modal properties if this decoherence condition is fulfilled.

The notion of decoherence to be used is the following. It is a general feature of the modal interpretation that if a system acquires a certain property, this happens by virtue of its interaction with the environment, as expressed in Eq. (11.1). As can be seen from this equation, in the interaction the system's property becomes correlated with a property of the environment. Decoherence is now defined to imply the irreversibility of this process of correlation formation: the rest of the universe retains a trace of the system's property, also at later times when the properties of the system itself may have changed. In other words, the rest of the universe acts as a memory of the properties the system has had; decoherence guarantees that this memory remains intact. For the state $|\Psi\rangle$ this means that in the Schrödinger picture it can be written in the following form:

$$|\Psi(t_n)\rangle = \sum_{i,j,\ldots,l} c_{i,j,\ldots,l} \, |\psi_{i,j,\ldots,l}\rangle \, |\Phi_{i,j,\ldots,l}\rangle, \tag{11.9}$$

where $|\psi_{i,j,\dots,l}\rangle$ is defined in the Hilbert space of the system, $|\Phi_{i,j,\dots,l}\rangle$ in the Hilbert space of the rest of the universe, and where $\langle\Phi_{i,j,\dots,l}|\Phi_{i',j',\dots,l'}\rangle = \delta_{ii'jj'\dots ll'}$. In (11.9) l refers to the properties $P_l(t_n)$, j to the properties $P_j(t_2)$, i to the properties $P_i(t_1)$, and so on.

The physical picture that motivates a $|\Psi\rangle$ of this form is that the final state results from consecutive measurement-like interactions, each of which is responsible for generating new properties. Suppose that in the first interaction with the environment the properties $|\alpha_i\rangle\langle\alpha_i|$ become definite: then the state obtains the form $\sum_i c_i|\alpha_i\rangle|E_i\rangle$, with $|E_i\rangle$ mutually orthogonal states of the environment. In a subsequent interaction, in which the properties $|\beta_j\rangle\langle\beta_j|$ become definite, and in which the environment 'remembers' the presence of the $|\alpha_i\rangle$, the state is transformed into $\sum_{i,j} c_i\langle\beta_j|\alpha_i\rangle|\beta_j\rangle|E_{i,j}\rangle$, with mutually orthogonal environment states $|E_{i,j}\rangle$. Continuation of this series of interactions eventually leads to Eq. (11.9), with in this case $|\psi_{i,j,\dots,l}\rangle = |\psi_l\rangle$.

If this picture of consecutive measurement-like interactions applies, it follows that in the Heisenberg picture we have $P_l(t_n)......P_j(t_2).P_i(t_1)|\Psi\rangle = c_{i,j,\dots,l}|\psi_{i,j,\dots,l}\rangle|\Phi_{i,j,\dots,l}\rangle$. Substituting this in the expression at the left-hand side of Eq. (11.8), and making use of the orthogonality properties of the states $|\Phi_{i,j,\dots l}\rangle$, we find immediately that the consistent histories decoherence condition (11.8) is satisfied. As a result, expression (11.7) yields a classical Kolmogorov probability distribution of the modal properties at several times.

11.7 Joint Probabilities of Events

In order to complete the spacetime picture that we discussed in section 11.5, we should specify the joint probability of events taking place in different spacetime regions. It is natural to consider, for this purpose, a generalization of expression (11.7). The first problem encountered in generalizing this expression to the relativistic context is that we no longer have absolute time available to order the sequence $P_i(t_1)$, $P_j(t_2)$, ..., $P_l(t_n)$. In Minkovski spacetime we only have the partial ordering $y < x$ (i.e., y is in the causal past of x) as an objective relation between spacetime points. However, we can still impose a linear ordering on the spacetime points in any region in spacetime by considering equivalence classes of points that all have space-like separation with respect to each other. The standard simultaneity hyperplanes provide examples of such classes. Of course, there are infinitely many ways of subdividing the region into such space-like collections of points. It will have to be shown that the joint probability distribution that we are going to construct is independent of the particular subdivision that is chosen.

Take one particular linear time ordering of the points in a closed region of Minkowski spacetime, for instance one generated by a set of simultaneity hyperplanes (i.e. hyperplanes that are Minkowski-orthogonal to a given time-like worldline). Let the time parameter t label thin slices of spacetime (approximating hyperplanes) in which small spacetime regions—'points'—with mutual space-like separation, are located. We can now write down a joint probability distribution for the properties on the various 'hyperplanes', in exactly the same form as in Eq. (11.7):

$$\begin{aligned} &Prob(P_i^*(t_1), P_j^*(t_2), ...P_l^*(t_n)) = \\ &\quad \langle\Psi|P_i^*(t_1).P_j^*(t_2).....P_l^*(t_n).P_l^*(t_n).....P_j^*(t_2).P_i^*(t_1)|\Psi\rangle. \end{aligned} \tag{11.10}$$

In this formula the projector $P_m^*(t_l)$ represents the properties of the spacetime 'points' on the 'hyperplane' labelled by t_l. That is:

$$P_m^*(t_l) = \prod_i P_{m_i}(x_i, t_l), \tag{11.11}$$

with $\{x_i\}$ the central positions of the point-like regions considered on the hyperplane. The index m stands for the set of indices $\{m_i\}$. Because all the considered point-like regions on the hyperplane t_l are space-like separated from each other, the associated projectors commute (the principle of microcausality). This important feature of local quantum physics guarantees that the product operator of Eq. (11.11) is again a projection operator, so that expression (11.10) can be treated in the same way as Eq. (11.7). In particular, we will need an additional condition to ensure that (11.10) will yield a Kolmogorovian probability.

The decoherence condition that we propose to use is the same as the one discussed in sections 11.5 and 11.6. Suppose that at spacetime point (x, t) the magnitude represented by the set of projector operators $\{P_k\}$ is definite-valued; P_l has value 1, say. In physical terms the notion of decoherence that we invoke is that in the course of the further evolution there subsists a trace of this property in the future lightcone of (x, t). That is, decoherence implies that on each space-like hyperplane intersecting the future lightcone of (x, t) there are (perhaps very many) local projectors that are strictly correlated to the earlier property P_l.

If this decoherence condition is fulfilled, we have because of the assumed permanence of the traces just as in section 11.6:

$$\langle\Psi|P_i^*(t_1).P_j^*(t_2).....P_l^*(t_n).P_{l'}^*(t_n)......P_{j'}^*(t_2).P_{i'}^*(t_1)|\Psi\rangle = 0, \quad i \neq i' \vee j \neq j' \vee .... \vee l \neq l'. \tag{11.12}$$

This makes (11.10) a consistent Kolmogorovian joint probability for the joint occurrence of the events represented by $P_i^*(t_1)$, $P_j^*(t_2)$, ...$P_l^*(t_n)$.

The projectors $P_i^*(t_1)$, $P_j^*(t_2)$, ...$P_l^*(t_n)$ depend for their definition on the chosen set of hyperplanes, labelled by t. Therefore (11.10) is not manifestly Lorentz invariant. However, the projectors $P^*(t)$ are products of projectors pertaining to the individual spacetime points lying on the t-hyperplanes, so (11.10) can alternatively be written in terms of these latter projectors. The specification of the joint probability of the values of a field at all considered 'points' in a given spacetime region requires (11.10) with projectors for all those points appearing in it. Depending on the way in which the spacetime region has been subdivided in space-like hyperplanes in the definition of $P_i^*(t_1)$, $P_j^*(t_2)$, ...$P_l^*(t_n)$, the projectors occur in different orders in this complete probability specification. However, there is a lot of conventionality in this ordering. All operators attached to point-like regions with space-like separation commute, so that their ordering can be arbitrarily changed. The only characteristic of the ordering that is invariant under all these allowed permutations is that if $y < x$ (i.e. y is in the causal past of x), $P(y)$ should appear before $P(x)$ in the expression for the joint probability. But this is exactly the characteristic that is common to all expressions that follow from writing out (11.10), starting from all different ways of ordering events with a time parameter t. All these expressions can therefore be transformed into each other by permutations of projectors belonging to spacetime points with space-like separation. The joint probability thus depends only on how the events in the spacetime region are ordered with respect to the Lorentz-invariant relation $<$; it is therefore Lorentz-invariant itself.

Within the just-discussed interpretational scheme, it is possible to speak of values of physical magnitudes attached to small spacetime regions. The notion of an event can therefore now be accommodated. This also gives us the conceptual tools needed to work with the notion of an object. An object can be treated as a particular distribution of field values. In particular, an (approximately) localized object can be regarded as a distribution of field values that vary continuously and fill a narrow world-tube in Minkowski spacetime. The notion of state localization to be used here implies that all observables have their vacuum expectation values in regions within the causal complement of the world-tube (Haag 1992, sect. V.5.3). Note that this does not mean that the probabilities of local observables are zero: the probabilities are the same as in the vacuum state. As we have noted in section 11.1, such local vacuum probabilities do not vanish.

11.8 A Possible Alternative: Perspectivalism

The approach sketched in the previous sections relied on the presence of decoherence mechanisms; only because we assumed that decoherence conditions were fulfilled could we obtain joint probabilities of events and Lorentz-invariance. Although this is perhaps enough from a pragmatic point of view, fundamentally speaking it seems undesirable that basic features like the existence of a joint probability distribution and Lorentz-invariance depend on the satisfaction of contingent, fact-like conditions. It seems therefore worth-while to consider the question of whether a more fundamental approach is possible that offers prospects of interpreting quantum field theory in a consistent realist way that is Lorentz-invariant from the outset. After all, that problems with joint probabilities and Lorentz-invariance can be made to disappear when decoherence mechanisms enter the stage is not really surprising. The interactions responsible for decoherence can be regarded as measurements performed by the environment, and it has often been observed in the literature that problems with joint probabilities and fundamental Lorentz-invariance do not show up in measurement results. By contrast, these topics cause considerable difficulty in the case of systems on which no external measurements are made—see Dickson and Clifton 1998 for an explanation of the problems in the context of modal interpretations.

The core of the problems identified by Dickson and Clifton (in Dickson and Clifton 1998) is that different Lorentz observers cannot—on the pain of inconsistency—use the same joint probability expression (11.2) for the simultaneous properties of two (more or less localized) systems in an Einstein-Podolsky-Rosen situation (see also Dieks 1998b). By exploiting the fact that in some Lorentz frames a measurement is made on system

1 before a measurement on system 2 takes place, whereas this order is reversed in other frames, Dickson and Clifton in essence show that the transitions undergone by the two systems during the measurements must be locally determined. Indeed, in the frames in which the measurements are not simultaneous there is no other measurement to take into account. But this result conflicts with the treatment given in a Lorentz frame in which the two measurements take place simultaneously, and in which such a local account is notoriously impossible.

One possible way out (see also Dieks 1998b) is to think of the properties assigned to the systems not as monadic predicates but as relations, in the manner explained in section 11.4. If, in the EPR-situation, the properties of systems thus need the specification of a perspective, it becomes natural to let this perspective play the role first played by the simultaneity relation in relating systems to each other. For example, one now needs to specify from which point of view the properties of system 1 are defined: from the point of view of system 2 before it has undergone a measurement, from the point of view of system 2 during the measurement, or from a still later perspective. Applying the rules of section 11.4, we find that different states are assigned to system 1, at one and the same spacetime point, from these different perspectives. If the properties of system 1 are regarded as non-relational, an immediate contradiction results because a system cannot simultaneously possess different states (reflecting its physical properties) at one spacetime point. Essentially, this is the contradiction derived by Dickson and Clifton. The contradiction disappears, however, if properties are relational. There is no logical difficulty involved in assuming that one and the same system, at one particular stage in its evolution, has different properties in relation to different reference systems.

If the object system and the reference system have definite positions in Minkowski spacetime, as in the above example, the object-perspective relation is associated with a spatio-temporal relation between two or more spacetime regions, which itself is Lorentz-invariant. If, moreover, all probability considerations are also made relative to perspectives, no problems with Lorentz-invariance can arise. It remains to be seen, however, whether this idea works also in more general situations.

11.9 Conclusion

We have discussed the possibility of interpreting quantum field theory in terms of a spacetime picture involving localized events, by means of the application of ideas from the modal interpretation of quantum mechanics. We have argued that the usual modal ideas, together with the fulfillment of a decoherence condition, ensure the existence of a simple, natural and Lorentz-invariant joint probability expression for the values of definite-valued observables at several spacetime locations.

Because of the uncountable infinity of degrees of freedom in the quantum field, the occurrence of decoherence, involving irreversibility, is something very natural to assume in the context of quantum field theory (Schroer 1999). Still, it would probably be more satisfactory to have a scheme in which the existence of a consistent joint probability distribution and Lorentz-invariance would not depend on such fact-like and contingent circumstances. We have therefore devoted a brief and tentative discussion to the idea of applying a new form of the modal interpretation, according to which properties have a relational character. This perspectival modal interpretation may offer prospects of obtaining a Lorentz-invariant picture even if no decoherence mechanisms are effective.

References

G. Bacciagaluppi and W.M. Dickson (1999): *Found. Phys.* 29: 1165.

G. Bene and D. Dieks (2001): A perspectival version of the modal interpretation of quantum mechanics and the origin of macroscopic behavior. E-print quant-ph/0112134. To be published in Foundations of Physics.

R. Clifton (1998): Beables in Algebraic Quantum Mechanics. In *From Physics to Philosophy: Essays in Honour of Michael Redhead.* Cambridge: Cambridge Univiversity Press.

R. Clifton (2000): *Phys. Lett. A* 271: 167.

W.M. Dickson and R. Clifton (1998): Lorentz-invariance in modal interpretations. In Dieks and Vermaas 1998.

D. Dieks (1989): *Found. Phys.* 19: 1397; *Phys. Lett. A* 142: 439.

D. Dieks (1998a): Preferred factorizations and consistent property attribution. In R.A. Healey and G. Hellman (eds.), *Quantum Measurement—Beyond Paradox.* Minneapolis: Univiversity of Minneapolis Press.

D. Dieks (1998b): Locality and Lorentz-covariance in the modal interpretation of quantum mechanics. In Dieks and Vermaas 1998.

D. Dieks (2000): *Phys. Lett. A* 265: 317.

D. Dieks and P.E. Vermaas, eds. (1998): *The Modal Interpretation of Quantum Mechanics.* Dordrecht: Kluwer Academic Publishers.

F. Dowker and A. Kent (1996): *J. Stat. Phys.* 82: 1575.

R. Griffiths (1984): *J. Stat. Phys.* 36: 219.

R. Haag (1992): *Local Quantum Physics.* Berlin: Springer-Verlag.

R. Healey (1989): *The Philosophy of Quantum Mechanics: An Interactive Interpretation.* Cambridge: Cambridge University Press.

G.C. Hegerfeldt (1974): *Phys. Rev. D* 10: 3320.

G.C. Hegerfeldt (1985): *Phys. Rev. Lett.* 54: 2395.

A. Kent (1995): *Phys. Lett. A* 196: 313.

D.B. Malament (1996): In defense of dogma: Why there cannot be a relativistic quantum mechanics of (localizable) particles. In R. Clifton (ed.), *Perspectives on Quantum Reality.* Dordrecht: Kluwer Academic Publishers.

C. Rovelli (1996): *Int. J. Theor. Phys.* 35: 1637.

B. Schroer (1999): Basic quantum theory and measurement from the viewpoint of local quantum physics. E-print quant-ph/9904072.

B. van Fraassen (1982): A modal interpretation of quantum mechanics. In E. Beltrametti and B. van Fraassen (eds.), *Current Issues in Quantum Logic.* New York: Plenum.

P.E. Vermaas and D. Dieks (1995): *Found. Phys.* 25: 145.

P.E. Vermaas (1999): *A Philosopher's Understanding of Quantum Mechanics.* Cambridge: Cambridge University Press.

Chapter 12

Measurement and Ontology: What Kind of Evidence Can We Have for Quantum Fields?

Brigitte Falkenburg
University of Dortmund

Abstract. *In the following, I deal with the ontology of quantum field theory (QFT) from a Kantian point of view, in terms of parts of empirical reality and their relations. In contradistinction to a formal ontology of QFT that is based primarily on the formal structure of the theory, I focus on the ways in which quantum fields can be measured, and on the structural features of empirical reality to which these measurements give rise. To approach the ontology of quantum fields in terms of measurement results in two paradoxes. First, ontology is about the structure of independent entities which belong to the furniture of the world, but measurements rely on interaction. Second, experimental evidence for quantum field theories is mainly based on particle tracks and other local phenomena. Thus, what kind of evidence can we have for the field structure of quantum fields? My paper attempts to unravel these paradoxes in the following steps. First, I give a rough sketch of the appearances of particle physics, the kinds of experimental evidence which count as tests of quantum electrodynamcs (QED) and the standard model of particle physics (1). In an intermezzo on Kant's view of scientific experience, I explain in which terms we might conceive of empirical reality beyond the claims of strict empiricism (2). Finally, I apply these ideas to the appearances of particle physics and suggest that they commit us to a relational ontology of QFT (3).*

12.1 QFT and the Appearances of Particle Physics

The appearance of a particle and the concept of a quantum field are roughly related as follows. The empirical appearance of a particle is a sequence of repeated position measurements in a well-defined spacetime region, that is, a track recorded by a particle detector during a certain short time interval. In data analysis, a collection of dynamical properties (physical magnitudes such as mass and charge) is attributed to any particle track. Operationally, a particle may thus be considered as a substance in Locke's sense, namely as a collection of empirical ideas or properties that go constantly together. This operational account of a particle may be related to Wigner's group theoretical definition of a particle. The Poincaré group describes the dynamics of particles or fields under spatio-temporal transformations. The irreducible representations of the Poincaré group are classified according to mass, spin and parity.[1] Thus a *particle* of given mass, spin and parity corresponds to an *irreducible representation* of the Poincaré group. In this way, a kind of particle is characterized by the invariant dynamic properties which are related to its spatio-temporal symmetries.

The definition can easily be generalized to the *charges* which are involved in the interactions of elementary particles. They are associated with the (local) dynamic symmetries of gauge theory. In this way, a kind of particle is characterized via the following collection of dynamical properties, which relate to the irreducible representations of the symmetry groups of the corresponding QFT:

$$| \, m, S, P, Q_k \, \rangle$$

Here m, S, P are mass, spin, and parity, and Q_k stands for a collection of charges which characterise the interactions of a specific kind of particle. According to the standard model of modern particle physics, these charges are: electric charge and flavor (electroweak interactions), and colour (strong interactions).[2]

The above "particle" definition is *very* general. It applies to any field, quantized or not, which is a solution of a relativistic field equation.[3] Obviously the definition deals with a *kind* of particle and not with the *indi-*

[1]Wigner 1939.

[2]Physics "beyond" the standard model introduces new charges, e.g. of supersymmetric interactions. Mass is the charge of gravity. Even without taking gravity into account, it has a special status in QFT (spontaneously broken symmetries and coupling to the Higgs field; the massive Higgs field quanta have not yet been measured).

[3]Its only restriction is: it applies to *free* fields. Here, "free" means "unbound", i.e. not part of a dynamically bound many-particle-system (such as the atom, or the solar system).

vidual collection of physical properties measured within a given spacetime region. The step from the quantum field, which describes kinds of particles and their interactions, to an individual particle requires *measurement.* This step is closely related to the fact that QFT, like any quantum theory, makes only probabilistic predictions about possible measurement results.

This fact gives rise to a crucial gap between quantum field theory and the appearances. QFT is crucially empirically underdetermined as regards the quantum field itself. The only quantities of QFT with an empirical content are the expectation values concerning subatomic transitions and particle interactions. Here we face the first one of the paradoxes mentioned above: If we ask for the ontology of QFT in terms of appearances, we face the results of interactions and measurements rather than the entities we have asked for.

Even the interactions of particles with quantum fields can not be resolved into processes which are related to well-defined magnitudes. In most cases the contributions of single Feynman diagrams to the perturbational expansion of a QFT can not be measured separately. Only superpositions of Feynman diagrams (or of corresponding "virtual particle" processes) are measured. However, the situation is strikingly different in the low energy and the high energy domain. In the low energy domain, there are several well-known high precision measurements which make it possible to single out the contribution of a single Feynman diagram to a quantum field, in lowest order of perturbation theory. In the high energy domain higher order contributions to the scattering amplitude become important. Here only *particle* tracks are measured, and we wonder how they relate to the quantum *field.* To explain the difference let us look at our best-known QFT, quantum electrodynamics (QED).

12.1.1 *The Low Energy Domain: Single Feynman Diagrams*

In the domain of low-energy phenomena it is at least possible to show *which* Feynman diagrams of the perturbational expansion of QED contribute *mainly* to the superpositions. This is shown especially by the best high precision tests of QED, the measurement of the Lamb shift in the hydrogen spectrum and the $(g-2)/2$-measurement of the electron or muon. Dirac theory alone predicts the wrong fine structure of the hydrogen spectrum (no splitting of the levels $S_{1/2}$ and $P_{1/2}$ for $n = 2$), and a gyromagnetic factor $g = 2$ of the electron or muon. Measurements show the Lamb shift of the hydrogen fine structure, and the anomalous magnetic moment of the electron. The anomalous difference $(g-2)/2$ between Dirac theory and actual magnetic moment can be measured with high precision from the spin precession of a charged particle in a homogeneous magnetic field. The next order QED correction stems from a single Feynman diagram which describes electron self-interaction. Here, theory and experiment agree to 7 digits, with a tiny discrepancy between theory and experiment in the 8^{th} digit. In such a case, the experiments are *almost* capable of singling out the "real" effect of a single Feynman diagram. Similarly in the case of Lamb shift. Here, the next order of perturbation theory gives a correction which is based on the Feynman diagrams for vacuum polarization and electron self-interaction. The correction shows that only 97% of the observed Lamb shift can be explained without the vacuum polarization term. A textbook of experimental particle physics tells us therefore that the missing 3% are "a clear demonstration of the actual existence of the vacuum polarization term".[4] As philosophers of QFT, we suspect that this is not *really* the case.

12.1.2 *The High Energy Domain: Scattering and Field Effects*

The basic appearances of QFT which I discuss in the following are the *particle tracks*, *scattering events* and *cross sections* of high energy physics, which have been measured in scattering experiments at big particle accelerators over the last decades. Particle tracks and scattering events are observed at the level of *individuals*, whereas the cross section is the *probabilistic magnitude* in which the notorious quantum measurement problem is hidden. QFT touches empirical grounds on both levels, as we shall see.

(i) Particle Tracks

At the level of individuals, a particle may be identified with a collection of physical properties which are once or repeatedly measured within a certain spacetime region. A particle track stems from repeated position measurements. The tracks are detected by sophisticated particle detectors: photographic plates with nuclear emulsions, bubble chamber tanks, drift chambers, "sandwich" scintillators, etc. In the measurement theory of particle physics, the *cause* of a particle track is explained as a kind of quantum

[4]Lohrmann 1992, p. 108 f.

process in which the detector detects repeatedly the *units of quantized magnitudes* such as charge or energy. In quantum field theory, such units of quantized magnitudes correspond to real (non-virtual) *field quanta* which are created and annihilated spontaneously due to field fluctuations. In a particle detector, such units are either absorbed, or they give rise to observable ionization processes.

The measured particle tracks are characterized by kinematic properties such as *energy and momentum* which obey relativistic kinematics, and by dynamic properties such as *mass, charge, spin* which determine the *particle type*: proton p, neutron n, electron e^+, positron e^-, muon μ, pion π, photon γ, etc.

Measurements of these properties are based on a semi-classical model of a particle trajectory. The model is empirically adequate at the probabilistic level, i.e. for *many* tracks which are generated under the same experimental conditions.[5] Charged particles are deflected by external electric and magnetic fields. The ratio q/m of charge and mass can be measured from the deflection, according to the classical Lorentz force law. If q and m are known, the momentum of a particle at the beginning of a curved track in a magnetic field can be reconstructed from the curvature. The reconstruction includes a detailed semi-classical calculation of the energy loss along the track, which is based on QED. In the low energy domain, non-relativistic quantum effects such as ionization loss are predominant, which give rise to smooth energy dissipation along a quasi-classical track. For very fast particles, QED processes such as *bremsstrahlung* and pair creation take place which result in "kinked" particle tracks. The kinks in the particle tracks indicate strong quantum fluctuations. In this way, specific QED processes such as *bremsstrahlung* and pair creation can be singled out at the level of individual appearances: they correspond to kinked tracks. At a probabilistic level, the relative number of such kinked tracks is in very good agreement with the QED predictions from the corresponding Feynman diagrams. Most often the shape of a given track even indicates whether *bremsstrahlung* or pair creation has occurred, at a kink. Here is one of the rare cases in which we can "see" individual contributions of a single Feynman diagram to a QFT process, with a high level of statistical confidence.

[5] Cf. Falkenburg 1996.

(ii) Scattering Events

A *scattering event* is reconstructed from 2, 3, 4,... particle tracks that start or stop within the same small spacetime region in a particle detector. The scattering events are interpreted as evidence of subatomic particle reactions in which "real" field quanta are annihilated and created. The measurements of the scattering events are based on *conservation laws* which relate to the *symmetries* of a QFT. The scattering events of high energy physics respect certain conservation laws which derive from the symmetries of the corresponding QFT. Vice versa, scattering events or particle reactions which have never been observed are assumed to violate conservation laws. They are related to hitherto unknown dynamic particle properties such as "strangeness" or "charm". In this way, the puzzling variety of particle tracks and scattering events found since the early years of accelerator physics had been resolved into the well-known classification of hadrons, which gave rise in the early 1960s to the quark model of particle physics and later to the standard model of particle physics.

Much more of QFT is needed to describe subatomic decays. The decays of unstable particles with a very short lifetime can not be directly observed but inferred from resonances occurring at a given scattering energy. The production of such a kind of particle, which is again characterized by mass, spin, parity and charges, causes a resonance effect in a measured cross section. That is, at a given scattering energy which corresponds to the mass of an unstable particle, an enormous increase of the relative frequency of particle reactions of a given type is observed. The Breit-Wigner formula describing a resonance derives from the S-matrix of a scattering process, and thus from QFT.

(iii) Cross Sections

The data analysis of high energy scattering experiments is based on the reconstruction of large samples of particle tracks and scattering events. Tracks and events are classified according to their kinematic and dynamic properties, and the so-called cross sections of several kinds of particle reactions are determined. The cross section is a *probabilistic magnitude* with the dimension of an area (unit: 1 barn = $10^{-24}cm^2$). It is defined as the relative number of scattering events with the same kinds of "output" particles or tracks with a certain 4-momentum, as compared to a given number of "input" particles or tracks. Operationally, a cross section is a measure of the *relative frequency* of a particle reaction of a given kind which corresponds to the *probability of a certain kind of measurement result.* All problems of the measurement process of a QFT are covered by this definition.

To relate the abstract formalism of QFT to experiment, the S-matrix for a certain kind of interaction has to be calculated from the coupled field equations for the corresponding fields. The S-matrix determines the theoretical expression for a measured cross section. The differential cross section $d\sigma/dq^2dE$ depends on the scattering energy E and the 4-momentum transfer q^2:

$$\left(\frac{d\sigma}{dq^2dE}\right)_{\mathrm{QFT}} \longleftrightarrow \left(\frac{d\sigma}{dq^2dE}\right)_{\mathrm{Exp}}$$

The whole theoretical machinery of QFT is involved in the calculation of the S-matrix. The starting point of any theoretical prediction of a measured cross section is the interaction term $\mathcal{L}'(x)$ of a Lagrangian of coupled fields. The lowest order of perturbation theory is the so-called Born approximation. It is based on the assumption that the incoming and outgoing particles, respectively the corresponding initial and final quantum field states,

$$|\, i\, \rangle = \lim_{t\to-\infty} |\, t\, \rangle$$

$$|\, f\, \rangle = \lim_{t\to+\infty} |\, t\, \rangle$$

are identical with free quantum fields. In the Born approximation, the interaction Lagrangian is given by the following expression which corresponds to the lowest order Feynman diagrams:

$$\langle\, f \mid S \mid i\, \rangle_{Born} = \delta_{fi} - i \int dt' \, \langle\, f \mid \int d^3x \mathcal{L}'(\mathbf{x}, t) |\, i\, \rangle$$

The S-matrix in the Born approximation is related to the differential cross section of a particle reaction as follows:

$$\left(\frac{d\sigma}{dq^2dE}\right)_{QFT} \sim f(p_1^i ... p_n^i; p_1^f ... p_m^f) \; |\langle\, f \mid S \mid i\, \rangle|^2$$

The magnitudes agree up to a flux factor which is not given here. The kinematical factor f depends on the momenta $p_1^i...p_n^i$; $p_1^f...p_m^f$ of the n incoming particles and m outgoing particles of the individual particle reactions.

(iv) Pointlike and Non-pointlike Structures

In high energy scattering experiments, the field structure of some particle is expressed in terms of non-pointlike structures. The semantics of these non-pointlike structures is expressed in terms of the equivalence of "scaling" and pointlikeness, and (vice versa) the equivalence of scaling violations and field-like distributions of quark or gluon momenta. (Here, "scaling" means that the cross section no longer depends on the scattering energy E of the experiment.) To interpret the data from such scattering experiments in terms of a non-pointlike particle or field structure depends crucially on two theoretical assumptions.

(i) The contributions of the interacting particles/fields to the Lagrangian $\mathcal{L}'(x)$ of the interaction are *separable.* Under this assumption (which corresponds to the Born approximation of scattering theory) we may use one kind of particles as "probe" particles which investigate the structure of another kind of particles functioning as scattering centers.

(ii) The cross section of the "probe" particles at some scattering center of arbitrary structure can unambiguously be interpreted in terms of deviations from the scattering at a "pointlike" scattering center. Under this assumption, it is possible to define form factors and structure functions which describe the dynamic structure of atoms and nucleons.[6]

These assumptions work for the atomic and nuclear form factors which have been measured in the 1950s, at low energies. In 1968, the scaling behaviour of scattering at higher energies confirmed the quark model of the nucleon. In the 1970s and 1980s, high precision measurements of structure functions that express the momentum distributions of the quarks within the nucleon

[6]The definition derives from the algebraic description of interacting quantum fields (current algebra). It gives rise to a relativistic generalization of the concept of a charge distribution. See Drell and Zachariasen 1961; Itsykson and Zuber 1985, pp. 159 f., 531 f.; and my discussion Falkenburg 1993, 1995.

were performed.[7] But with increasing energy, scaling violations have been observed. According to quantum chromodynamics (QCD), they indicate an increasing amount of quark-antiquark pairs and gluons due to interactions of the quark fields within the nucleon.[8] Thus the dynamic structures measured in scattering experiments depend crucially on the energy of the scattered particles. The larger the scattering energy is, the more structure is observed, in accordance with QFT predictions of pair creation and other perturbative processes beyond the Born approximation. Are these structures *observed* at a given scattering energy, or are they *generated* by the experiment? Here, a radically new kind of context dependence enters our concept of physical objects. The dynamic structure of an object depends crucially on the kind of measurement we perform. Thus in a certain sense, the usual "correspondence" criteria for the truth or falsity of our theoretical models no longer apply. In my view this is a case for a *relational ontology*, as I hope to explain in the last part of my paper.

12.2 Intermezzo: A Kantian Account of Ontology

According to Kant, empirical science investigates the structure of a world of appearences, that is, of empirical reality, and not a world of things in themselves which corresponds to the claims of metaphysical realism. From a Kantian point of view, it is not only paradoxical but indeed contradictory to talk about *independent* entities which are at the same time considered to be *objects of our knowledge*.[9] The first assumption gives to an entity the status of a thing-in-itself which does not belong to empirical reality, whereas according to the second one an entity is given by means of experience and measurement.

Why do I propose a Kantian account of ontology? Traditionally, ontology has been understood as a doctrine of being-in-itself (Aristotle), or of independent entities such as Leibnizean monads. Traditional ontology aims at giving a true description of the real world. Kant criticised traditional ontology because it embraces objects such as God or the world as a whole

[7] Perkins 1987, pp. 262 f., explains the experimental results and their interpretation in terms of the quark-parton-model. Riordan 1987 tells the history.

[8] Perkins 1987, pp. 302 f.

[9] This is the rationale of Kant's doctrine of the antinomy of pure reason. For a detailed analysis, cf. Falkenburg 2000, pp. 177 ff.

to which we have no epistemic access. As a consequence, he attributed objective reality only to the objects of possible experience. In Kant's view, an ontology which is in accordance with the conditions of the possibility of objective knowledge is restricted to a doctrine of the structure of the empirical world. Such an ontology is defined in terms of *our cognitive capacities.* In the exact sciences, these cognitive capacities give rise to elaborated use of the experimental method, and of mathematics as part of the data analysis of experimental results.

A Kantian account of ontology is in accordance with *empirical realism,* or, as Putnam puts it, *internal realism.* In my view, such a modest, empirical ontology of the phenomenal world is much more appealing than a defense of metaphysical realism and the associated God's eye view of the world which Putnam criticizes in a Kantian line of reasoning. Any ontology of physical theory which does *not* rely on epistemological considerations is in danger of falling back into some sort of pre-Kantian dogmatic metaphysics. How can we avoid this danger in the case of QFT?

Let me start with a sketch of Kant's view of empirical reality. In contradistinction to traditional or 20th century empiricism, Kant emphasizes that experience depends on theory. For the purposes of the present paper, it suffices to rely on the following basic features of his view of empirical reality. Empirical reality is a *relational structure.* It is given in terms of relations which hold between appearances. All appearances are relational entities, too. No appearance has internal properties in a Leibnizean sense.[10] In other words, Kant's empirical reality is a mereological sum of relational parts without least parts. His ontology of appearances is a mereology without atoms. In contradistinction to the appearances, things in themselves are relationless, like Leibnizean monads. Since relationless entities can neither be experienced nor measured, they do not belong to empirical reality.

However, empirical reality embraces much more structure than the mere empirical relations that hold between appearances. Kant's empirical realism admits several kinds of parts of reality or objects of possible experience. The first is straightforward, the second invokes the postulates of empirical thinking, the third admits experimental data.

1. The basic ingredients of empirical reality are *appearances*, that is,

[10]Kant says: "The inner determinations of a substantial phenomenon in space [...] are nothing but relations, and it is itself entirely a sum total [Inbegriff] of mere relations." Kant 1787, p. 321 = 1781, p. 265.

spatio-temporal objects or events which are immediately perceived and related to each other according to the principles of pure understanding.

2. In addition, Kant accepts *causes of appearances*, that is, things connected with appearances according to the three analogies of experience. According to the analogies of experience, the relations between all parts of empirical reality are based on *a priori* principles of conservation of substance, causality, and universality of interaction. We may understand these principles as some kinds of *a priori* guides to inferences to the best explanation. (Whether they commit us to a classical, deterministic ontology or not is an open question of Kant interpretation which I cannot discuss here.) Kant's own example of the existence of an unobservable part of empirical reality is "the existence of magnetic matter penetrating all bodies" which we infer "from the perception of attracted iron filings, although an immediate perception of this matter is impossible for us given the constitution of our organs".[11]

3. Kant knew very well that the objects of empirical science are neither immediately perceived nor simply related to perceptions. He was aware that empirical science is based on appearances which are obtained by means of experimental investigation and measurement. From Kant's point of view, experiments are theory-laden in quite another way than any kind of nonscientific experience. In his view, an experiment is a specific question which we put to a specific part of nature (or empirical reality) — "like an appointed judge who compels witnesses to answer the questions he puts to them."[12] Such a question to nature is put under certain further *a priori* presuppositions, which add to the general *a priori* principles of the pure understanding. They concern our measurement devices, and the processes which we investigate in experiments. (We might call the first kind of *a priori* 'absolute'and the second kind 'relative', in the spirit of Reichenbach's distinction between Kant's own *a priori* of space and time, and the *a priori* of our physical assumptions about spacetime.)

Kant explained the relation between physical theory and observation in terms of a hypothetical-deductive approach which is strikingly modern. Even though he wanted to explain all physical objects in terms of Euclidean spacetime, Newtonian forces, relational substances, and an ether-like matter which penetrates all bodies, his relational view of empirical reality is

[11]Kant 1787, p. 273 = 1781, p. 226 (chapter on the "postulates of empirical thinking").

[12]Kant 1787, p. XIII.

quite liberal. It can cope with the modern concepts of field and interaction as well as with complicated ways of tracing back from experimental data to theoretical explanations. Let us summarize his view of empirical reality as follows:

I. Empirical reality is a relational structure. All of its parts, appearances as well as unobservables, are relational entities. Nothing in empirical reality has internal properties in a Leibnizean sense.
II. Empirical reality is given in terms of relations which hold between
   1. Appearances, and/or
   2. Unobservable entities connected to the appearances according to the three analogies of experience, and/or
   3. Experimental data resulting from a specific question put to nature under certain conditions.

We should be aware that it is only meaningful to talk about *parts* of empirical reality if we assume that we are able to *separate* such parts somehow, at least on the basis of well-confirmed theoretical principles. Unrestricted separability of empirical reality into parts is a further *a priori* assumption of Kant's theory of nature. Needless to say that it is closely related to the Galileian resolutive-compositive method of empirical science. (In Kant's view, the separability condition is related to the axioms of pure intuition, that is to the *a priori* principle that all appearances which are given in pure intuition are extensive magnitudes. Today, we should be more liberal and admit also more abstract, non-spatial part-whole relations between the appearances.)

12.3 Toward an Empirical Ontology of QFT

With these Kantian *a priori* assumptions in mind, let me review the kinds of evidence we have in the domain of QFT. Which ontological claims about substance, causality, and interaction do they support? In particular, I am looking for QFT-specific answers to the following two questions:

I What kind of experiments can we perform to test the specific structure of a QFT?

II What kind of cause can we infer from the appearances of QFT? (To which specific causal assumptions are we committed by the experimental data, according to Kant's three analogies of experience?)

The first question aims at a *complete account of the conditions of possible experience*, in the domain of QFT. In the first part of my paper, I discussed only certain types of experiment. But doing so, I started with the usual distinction of "low energy" and "high energy" experiments. The distinction covers the whole *energy scale* of QFT phenomena. Heuristically, energy is a measure of length. The larger the energy at which a scattering experiment is performed, that is, the higher the energy of the particle beam coming from the accelerator, the smaller subatomic structures can be resolved in the subsequent data analysis of the resulting particle tracks (given that the statistics of the experiment is good enough, i.e. thousands of particle tracks have to be analyzed). The underlying law is analogous to the formula for the resolution of an X-ray microscope,[13] and it derives from the de Broglie relation between momentum and wavelength. In the sloppy language of high energy physics, this heuristic relation is expressed as follows. The higher the energy of an interaction, the smaller the associated length scale. In this sense, by studying the typical low and high energy QFT tests, one has access to the whole empirical domain of QFT and to the ways in which its parts are accessible by experiments.

The second question aims at reconstructing these QFT-specific conditions of possible experience in terms of Kant's categories of pure reason. Regarding the term "cause" we should be liberal. We should also admit probabilistic explanations, and *explain* an individual scattering event only *after* it has already happened, that is, after the corresponding particle tracks have already been measured. In doing so, we avoid the burden of the quantum measurement problem.

12.3.1 *Low Energy Phenomena*

In the low energy domain, it is possible to defend the claim that "there are" individual appearances which correspond to terms of a QFT such as the next-to-lowest-order Feynman diagrams of QED. Recall the Lamb shift of hydrogen and the anomalous magnetic moment of the electron. The Lamb

[13] Cf. Falkenburg 1993, 1995 pp. 140 ff.

shift as well as the measured $(g-2)/2$-value are due to parts of empirical reality which are represented by the corresponding Feynman diagrams. But what kinds of entities are the parts of empirical reality "behind" the measurement results? We may say that the Lamb shift is caused by the interaction of a jumping electron with its own radiation field. We observe radiative transitions which indicate that the corresponding energy level of the hydrogen atom is shifted. The appearance is theory-dependent, it is the difference between our theoretical expectation (from the Dirac equation) and the observed radiative transitions. Obviously, we should not understand this statement *exactly* in the sense of Kant's principle of causality. The experiment does not single out a *cause* (second analogy), it singles out a contribution (or to be more precise, a superposition of two contributions) to an *interaction* (third analogy). In contradistinction to this phenomenon, the anomalous magnetic moment of the electron has the character of a persistent physical property. In Kantian terms, it is an intensive magnitude. The exact intensity of this magnitude is due to a certain kind of interaction which is permanently taking place in the investigated physical system and which is singled out by the measurement.

I conclude that in the low energy domain of QFT it is possible to separate parts of empirical reality with persistent properties. Sometimes, these properties look like non-relational properties (such as the anomalous magnetic moment). But indeed all of them are due to *interactions.*

12.3.2 *High Energy Phenomena*

In the high energy domain, there are two kinds of appearances: (1) particle tracks, scattering events and cross sections measured in high energy scattering experiments; and (2) data from astrophysics, such as information about radiations from supernovae, or inferences to the early "hot" phases of the universe. In astro-particle physics where high energy cosmic rays are detected, both domains overlap. I focus on my earlier discussion of scattering experiments. Their results seem to restrict us to a probabilistic ontology of quantum fields. What can we make out of them from a Kantian point of view? Do we gain grounds beyond empiricist positions such as van Fraassen's constructive empiricism or Suppes' probabilistic metaphysics? It is indeed possible to make the ontology "behind" the phenomena of QFT much more precise, in terms of Kant's postulates of empirical thinking. How do they apply to (i) particle tracks, (ii) scattering events, and (iii) measured cross sections?

Wigner's particle definition (i.e. the irreducible representations of the Poincaré group) corresponds to non-interacting fields, that is, to non-relational entities. Their dynamic properties (mass, spin, parity, different charges) are Leibnizean, monadic properties. According to Kant, such properties do not belong to empirical reality. In perfect accordance with this claim, unrenormalized free quantum fields do not belong to the empirical ontology of QFT. To give them physical meaning, we have to renormalize mass and charge. If we want to describe them before and after interactions, we consider them as *asymptotically* free. These observations give further support to my remarks concerning the low energy domain. For "physical" quantum fields and/or field quanta, the maximum of ontological independence is *asymptotic freedom*. If and only if the quantum states involved in a scattering process are asymptotically free, it makes sense to calculate the S-matrix of the scattering in Born approximation, and to add radiative corrections.

For asymptotically free states, Kant's empirical postulates fit in with a relational ontology of quantum fields which is based on the following claim:

(FQ) Field quanta are quantized units of physical magnitudes which are exchanged in interactions and measured by means of particle detectors.

From an ontological point of view, however, we want to single out the *relata* of these interactions, that is, the *causes* of the observed particle tracks and scattering events. What can we tell about them? The particle tracks in a detector and the scattering events reconstructed from such tracks are direct evidence for typical QFT processes. They are *caused* by discrete field quanta which according to Kant's criteria belong to empirical reality.

(i) Particle Tracks

The discrete field quanta "are" what we measure. Any event in a particle detector is due to an exchange of a unit of a quantized magnitude. Any particle track stems from a sequence of position measurements. QFT tells that each quantum which results an observable position measurement corresponds to a definite change of state of a quantized field. The kink in a particle track indicates that a specific QED process such as *bremsstrahlung* or pair creation has taken place. The inference to such a process explains what *has* happened at a certain time in empirical reality. The corresponding (single) Feynman diagram explains a concrete appearance of particle physics, in accordance with Kant's principle of causality and postulates of empirical thinking.

(ii) Scattering Events

Scattering events which give evidence of the creation and/or annihilation of particles also show that field quanta belong to empirical reality. Scattering events are the fingerprints of individual subatomic interactions. They show again that Kant's principle of causality applies to the domain of QFT, and that the resulting ontology of QFT is relational. But field quanta do not count as substances in Kant's sense, that is as magnitudes which are permanently conserved. Field quanta show only up during subatomic interactions. They are "ephemerals with a particle grin".[14] Besides the mass and electromagnetic structure of the particle detector only one kind of substance is involved in a scattering event, namely the specific conserved magnitudes of a given scattering process or particle reaction. These conserved magnitudes correspond to the marks of Wigner's particle definition, mass, spin, parity, and different charges. There are processes, however, in which these magnitudes are not conserved. Mass may change according to the relativistic mass-energy equivalence. In addition, electroweak interactions violate parity as well as CP conservation. The respective experimental results indicate that QFT is about parts of empirical reality where *substance vanishes*, not only in Aristotle's sense of individuals (*no* part of a quantum system is an individual!), but also in Locke's sense of simple ideas that go constantly together, or in Kant's sense of properties which persist in empirical change. In the domain of QFT the number of valid conservation laws decreases. CPT invariance and energy conservation are (presumably) the only principles left over.

Particle tracks and scattering events are *individual appearances.* Particle tracks indicate empirically that there are individual field quanta which belong to asymptotically free fields, and which have interactions indicated by scattering events. In accordance with Kant's postulates of empirical thinking, philosophers may follow the physicists' inference that the QFT-entity which interacts with the detector is a quantum field. Such an inference is analogous to Kant's claim that the attraction of the iron filings is due to

[14]Redhead 1982, p. 83.

some kind of magnetic matter. But what kind of entity is a quantum field? For reasons which are made explicit in other contributions to this volume, we can neither identify the quantum field with a classical field, nor with a collection of field quanta.

QFT gives a formal description of quantum fields in terms of field operators and quantization rules. Due to the quantization rules, these field operators do not have any obvious meaning beyond the usual operational, probabilistic interpretation of a quantum theory, that is, beyond the appearances described above. In addition to (FQ), however, the uninterpreted parts of QFT, i.e. the Feynman diagrams, tell us several kinds of causal stories about the kinds of processes which may give rise to field quanta. QFT explains that field quanta may arise from the interactions of a quantum field with (i) another quantum field, or (ii) itself, or (iii) the vacuum. What about these uninterpreted parts of QFT? There are only two options. Either they are fictitious, or they talk about unobservables which cause the appearances in the sense of Kant's postulates of empirical thinking. In the latter case, they are about *some* parts of independent reality which do *not reduce* to the appearances but *cause* them. To be just fictitious, QFT is far too splendid. (Remember the high precision low energy tests of QED.) Thus Kant's postulates of empirical thinking suggest that *there are* quantum fields which reduce neither to units of quantities measured in a particle detector nor to collections of field quanta.

However, they are not on a par with Kant's magnetic matter which penetrates all bodies. They are *no substances.* In any interaction, they change state. They may annihilate to *almost nothing*, that is, to a vacuum filled with energy which gives rise to particle creation. They do not have well-defined amplitudes. But what if they are none of these kinds of entities? Let us look at the informal answer a physicist might give us:

(QF) A quantum field is a non-local dynamic structure with local interactions.

Due to the quantum measurement problem, we have to understand this answer in probabilistic terms. Otherwise, the ontology of a quantum field would collapse into that of a classical field. But together with the usual probabilistic interpretation of a quantum theory, (QF) does not give us any ontology beyond the laws of QFT itself. Is there a more fruitful use of Kant's postulates of empirical thinking, concerning quantum fields? Is it possible to avoid the dilemma of operationism here and formalism there, that is, of identifying the cause "quantum field" either with its observable effect "measured field quanta", or with the mere formal structure of QFT itself? From a relational point of view, the answer is "yes". To confirm this view, let us look again at the measured cross sections corresponding to the S-matrix elements of QFT.

(iii) Cross Sections

The measured cross section of a certain kind of particle reaction is a *probabilistic* magnitude. Its probabilistic cause (in the sense of Kant's postulates of empirical thinking) is the corresponding S-matrix element, respectively its counterpart in empirical reality. The S-matrix is calculated from an interaction Lagrangian $\mathcal{L}'(x)$ of a QFT. S-matrix elements as well as the corresponding $\mathcal{L}'(x)$ describe the interaction of coupled quantum fields. In the Born approximation both correspond most often to single Feynman diagrams.[15] Again, this fact is only in accordance with a relational account of empirical reality. The formal description of a single quantum field on its own has no counterpart in empirical reality. It belongs neither to the appearances nor to their unobservable causes. Again, I emphasize that the interaction of a quantum field with (i) itself, or (ii) another field, or (iii) the vacuum, is described by a certain type of Feynman diagram which is only a *symbolic* part of the formal perturbational expansion of an interaction. Only in *some* cases, it may be singled out at least *approximately*.

(iv) Pointlike and Non-pointlike Structures

The pointlike or non-pointlike structures measured in high energy scattering experiments are most tricky. In my view, they show that in the domain of QFT, the structure of the appearances (or of empirical reality) depends crucially on our theoretical and technological abilities to separate the partners of scattering processes. If we can calculate a scattering matrix that factorizes nicely into parts corresponding to the various kinds of scattered particles, we have a comprehensible model of what is going on in a scattering process. If we can perform a scattering experiment that realizes our theoretical model approximately, we know what is going on in empirical reality at a given energy scale. As far as we are able to do so, our QFT models

[15] An important exception is Bhabha scattering, $e^+e^- \to e^+e^-$. Here, the Born approximation includes a superposition of "spacelike" and "timelike" processes.

of scattering processes represent field structures which make up parts of empirical reality. But this empirical reality is *not* made up of any *independent* furniture of the world (except, presumably, energy as a truly conserved quantity). Its structure is relational and context-dependent. It underlies the appearances which we generate in the scattering experiments of high energy physics. Presumably it underlies also processes at the same energy scale in the early universe. The higher the scattering energy, however, the more the relational parts of empirical reality become entangled.

12.4 Conclusions

To sum up, according to a Kantian view of empirical reality as well as experimental evidence, the (empirical) ontology of QFT is relational. As parts of empirical reality, quantum fields have no reality "on their own". Vice versa, on their own they do not belong to empirical reality. Under the specific empirical conditions of high energy physics, they exist only as strongly coupled entities. In the low-energy domain (which corresponds to relatively large parts of empirical reality), they may count as asymptotically free. Whenever we use quantum fields or field quanta to explain the appearances of particle physics we talk about subatomic interactions, about dynamic processes which happen at a very small scale. Whenever we are able to disentangle a quantum process into well-defined contributions of units of quantized magnitudes and to associate these magnitudes with point-like structures, we may talk in a sloppy way about a particle. If not, we may talk in the same sloppy way about a field. Both ways of talking neither mean classical particles nor classical fields, nor do they mean independent entities at all. Whenever we talk about free quantum fields, we talk in abstract and symbolic language (to adopt Niels Bohr's words), without referring to any concrete part of empirical reality. Strictly speaking, quantum fields on their own have no empirical reality. From a Kantian point of view, there is thus no need for an ontology of QFT in the traditional sense of a doctrine of entities-as-such. On the other hand, Kant's relational account of empirical reality helps us to escape from empiricism or operationism. Kant's most *famous* theoretical insight is that experience is based on *a priori* judgments. When we ask for an empirical ontology of QFT, however, his most *important* a priori insight is that empirical reality is *relational*. His relationalism is at odds with any attempt to explain the appearances in terms of entities on their own, and so is the structure of the appearances of QFT.

References

Drell, S.D., and Zachariasen, F. (1961): *Electromagnetic Structure of Nucleons.* Oxford.

Falkenburg, B. (1993): The concept of spatial structure in microphysics. *Philosophia Naturalis*, 30:208.

Falkenburg, B. (1995): *Teilchenmetaphysik.* Heidelberg: Spektrum Akademischer Verlag.

Falkenburg, B. (1996): The analysis of particle tracks: a case for trust in the unity of physics. *Stud. Hist. Phil. Mod. Phys.* 27:337–371.

Falkenburg, B. (2000): *Kants Kosmologie*, Frankfurt am Main: Klostermann.

Itzykson, C., and Zuber, J.-B. (1985): *Quantum Field Theory* (3rd edition). New York: McGraw-Hill.

Kant, I. (1781 = ed. A) / (1787 = ed. B): *Critique of Pure Reason.* Transl. by P. Guyer and A. Wood, Cambridge: Cambridge University Press, 1998.

Lohrmann, E. (1992): *Hochenergiephysik.* Stuttgart: Teubner.

Perkins, D.H. (1987): *Introduction to High Energy Physics* (3rd edition). Menio Park, California: Addison-Wesley.

Redhead, M. (1982): Quantum field theory for philosophers. In: P. D. Asquith and T. Nickles (eds.), PSA 2:83.

Riordan, M. (1987): *The Hunting of the Quark.* New York: Simon & Schuster.

Wigner, E. (1939): On unitary representations of the inhomogeneous Lorentz group. *Annals of Mathematics*, 40:149–204.

Chapter 13

Renormalization and the Disunity of Science

Nick Huggett
University of Illinois at Chicago

13.1 The Need for Renormalization

The kind of physical quantity that one wants to calculate in quantum field theory (QFT) is the probability for colliding particles with given momenta to scatter to produce other particles with other given momenta. Indeed it is very natural to identify the complete set of amplitudes of a theory with the theory itself, since this set exhausts the physical content of the theory, and since it picks out something definite, independent of any reference to perturbation theory or renormalization. I will primarily use the term in this sense, though it perhaps obscures questions of ontology beyond the phenomena.

In QFT, as in most realistic examples in quantum mechanics (QM), amplitudes cannot be found as exact functions, but rather via perturbation theory: one expands in powers of a small perturbative parameter—typically the interaction strength λ—about a known solution—the free field in QFT. Schematically,

$$\Gamma^{(N)}(p_1, \ldots, p_N) = \Gamma_0 + \lambda \cdot \Gamma_1 + \lambda^2 \cdot \Gamma_2 + \ldots, \qquad (13.1)$$

where $\Gamma^{(N)}$ is the amplitude for a process involving N particles with incoming and out-going momenta $p_1, \ldots, p_N$, and $\lambda^n \cdot \Gamma_n(p_1, \ldots, p_N)$ is the n^{th} order contribution to the process.

Unfortunately, in most cases, far from having a finite sum, the individual terms of the expansion are divergent. For instance, in $\lambda\varphi^4$ theory (in four spacetime dimensions), the leading corrections are quadratically and logarithmically divergent (higher terms contain dependencies on these terms, and so also diverge):

$$\Gamma(p) \sim \Gamma_0 \ + \ \lambda \cdot \int_0^{\infty} \mathrm{d}^4q/q^2 \ + \ \lambda^2 \cdot \int_0^{\infty} \mathrm{d}^8q/q^8 \ + \ \ldots \tag{13.2}$$

To quote from Teller (1995, 150), "Occurrence of infinities, one would think, should make it clear that some fatal flaw infects the fundamental theory, the approximation scheme, or the system of the two taken together." Now as Teller shows, standard techniques of renormalization enable one to manoeuver around this problem, in a formally legitimate way. But in this paper I will argue that, *prima facie* questions of the significance—as opposed to the consistency—of those techniques, raise important doubts about the standing of QFT: in particular, whether QFTs can be taken seriously as candidate true theories. What I aim to explain is how the 'renormalization group' (RG) provides a map that helps us understand the infinities and the manoeuvering, and hence removes those doubts.

13.2 The Method of Renormalization

Recall, in its broadest outlines, what it means to renormalize (I will assume familiarity with such standard presentations as Ryder, 1985, chap. 9, or Cheng and Li, 1984, chap.2). First we make our working theory—or at least the individual terms in its expansion, since convergence is a separate issue— finite by introducing a cut-off, $\bar{\Lambda}$ in the upper limit of momentum integration.

$$\Gamma(p, \bar{\Lambda}) \sim \Gamma_0 \ + \ \lambda \cdot \int_0^{\bar{\Lambda}} \mathrm{d}^4q/q^2 \ + \ \lambda^2 \cdot \int_0^{\bar{\Lambda}} \mathrm{d}^8q/q^8 \ + \ \ldots \tag{13.3}$$

This expansion still describes a QFT, but one for which momentum has an upper bound—equivalently for which length has a lower bound.

However, there is no natural cut-off—or rather, there is no known physical effect that might determine such a cut-off—so the cut-off theory is underdetermined. Thus we 'renormalize' the theory's mass and charge parameters and field strength, replacing the measured physical values with 'bare parameters' designed to cancel the divergent cut-off dependence in the $\bar{\Lambda} \to \infty$ limit. Naturally, these parameters must themselves be functions of the cut-off if they are to achieve this end. Schematically:

$$\text{physical mass} \quad m \quad \longrightarrow \quad m_0 \equiv m - \Sigma(\bar{\Lambda}, \mu) \quad \text{bare mass}$$

$$\begin{array}{lcll} \text{physical charge } \lambda & \to & \lambda_0 \equiv \lambda - \Gamma(\bar{\Lambda}, \mu) & \text{bare charge} \\ \text{physical field } \varphi & \to & \varphi_0 \equiv Z_\varphi(\bar{\Lambda}, \mu) \cdot \varphi & \text{bare field} \end{array} \quad (13.4)$$

for some (specified) functions Σ, Γ and field rescaling factor Z_φ. Note that defining the bare quantities requires the introduction of an arbitrary mass scale μ.

The cut-off theory obtained by making these substitutions is the 'bare theory', and if done correctly—mechanical methods exist for doing so—then any divergent terms will be canceled. In particular, a 'renormalized' theory, depending on the physical, 'renormalized parameters', will be related to the bare theory according to:

$$\Gamma_r^{(N)}(p, m, \lambda, \mu, \bar{\Lambda}) = Z_\varphi^{-N/2}(m_0, \lambda_0, \mu, \bar{\Lambda}) \cdot \Gamma_0^{(N)}(p, m_0, \lambda_0, \bar{\Lambda}). \quad (13.5)$$

The crucial things to absorb from this kind of equation are (a) the fact that it is a relation between the physical content of two theories, the amplitudes Γ_0 of a bare theory and the amplitudes Γ_r of a renormalized theory, and (b) how dependence on various variables enters the relation. This latter point is crucial, since the RG considers how physical quantities vary with respect to such variables, particularly p, μ and $\bar{\Lambda}$.

Finally, in the limit $\bar{\Lambda} \to \infty$, assuming the renormalization was indeed done correctly, any divergent cut-off dependence is canceled into the parameters of the bare theory, and we are left with a finite renormalized theory, independent of any cut-off—a continuum theory, with no finite minimum length.

$$\Gamma_r^{(N)}(p, m, \lambda, \mu) = \lim_{\bar{\Lambda} \to \infty} Z_\varphi^{-N/2}(m_0, \lambda_0, \mu, \bar{\Lambda}) \cdot \Gamma_0^{(N)}(p, m_0, \lambda_0, \bar{\Lambda}). \quad (13.6)$$

Seemingly miraculously, this theory gives just the right physical answers, once the physical mass and charge, m and λ, are measured.

One point is most important to recognize: in the process of renormalization, a new quantity with dimensions of mass—i.e., μ—has been introduced, on which the amplitude depends. But this is just an artifact of the calculation, forced by dimensional considerations, but with no apparent physical significance.

13.3 The Renormalization Group

Now, how are we to understand renormalization? Work (especially) since the mid-1970's on the RG provides a powerful picture of what is going on (we will follow Le Bellac, 1991, chap. 7). Interestingly, the significant contributions to our understanding of this topic have come both from field theory (e.g., Gell-Mann and Low, 1954) and from the study of second order phase transitions in solid state physics (e.g., Wilson and Kogut, 1974). This interplay is due to a deep (and perhaps surprising) formal analogy between path integral QM and statistical mechanics in general, and between lattice QFT and solid state physics in particular. Fascinating though this topic is, for simplicity here we will focus on the RG solely within QFT (see Wilson and Kogut, 1974, sec. 10, or Shenker, 1984, for more on the analogy).

Consider again the relation between bare and renormalized theories, equation 13.5. We will, however, make one important change; we will write it in terms of 'dimensionless' parameters. That is, in $\hbar = c = 1$ units every physical quantity can be expressed as a power of mass: e.g., momentum is expressed as a mass, and length as a single inverse power of mass. In general, we write the mass dimensions of a parameter λ as $[\lambda]$. Then, given a natural mass scale in a theory—e.g., μ—one can introduce dimensionless parameters g, according to $\lambda \equiv g \cdot \mu^{[\lambda]}$. In all that follows we will be implementing the RG in terms of the dimensionless parameters, as this offers the clearest view of its meaning.

So rewrite equation 13.5, using g to denote all the (dimensionless) parameters, including mass, and let the cut-off be Λ:

$$\Gamma_r^{(N)}(p, g, \mu, \Lambda) = Z_\varphi^{-N/2}(g_0, \Lambda/\mu) \cdot \Gamma_0^{(N)}(p, g_0, \Lambda). \tag{13.7}$$

This equation appears to be a purely formal relation—the penultimate step in our renormalization procedure before we take the cut-off to infinity—but the RG gives it a powerful interpretation.

First, it is central to renormalization that in the large cut-off limit, the renormalized amplitude is cut-off independent; that's what makes the renormalized theory finite in the limit. Therefore the total derivative with respect to Λ of either side of equation 13.7 vanishes in that limit. Setting the RHS to zero and manipulating yields:[1]

$$\left[\frac{\partial}{\partial \ln \Lambda} + \beta(g_0) \cdot \frac{\partial}{\partial g_o} - \frac{N}{2} \cdot \gamma(g_0)\right] \cdot \Gamma_0^{(N)}(p, g_0, \Lambda) = 0, \qquad (13.8)$$

the 'Callan-Symanzik' (CS) equation, a flow equation for the bare theory, with

$$\gamma(g_0) = \frac{\mathrm{d} \ln Z_{\varphi}^{-N/2}}{\mathrm{d} \ln \Lambda} \quad \text{and} \quad \beta(g_0) = \frac{\mathrm{d} g_0}{\mathrm{d} \ln \Lambda}. \qquad (13.9)$$

We obtained equation 13.8 from the fact that the renormalized theory is independent of the cut-off. The CS equation then describes how the dimensionless bare parameters and the field rescaling must change to compensate for the effect on the bare theory of varying the cut-off. Let's spell that out a little.

For some fixed renormalized QFT, we have a bare theory with cut-off Λ, which, up to field rescaling, has the same physical consequences: precisely, the bare and renormalized theories agree on all scattering amplitudes. But Λ parameterizes a family of bare theories, each of which by equation 13.7 is equivalent to the same renormalized theory. And the only way that these theories, which after all have different cut-offs, can correspond to the same physics is if they also have different values for their parameters. The β- functions describe how the parameters must vary with Λ to maintain equivalence.

There's a clear physical interpretation of the matter: a maximum momentum—the cut-off—is equivalent to a minimum distance, so one can speak of a cut-off QFT as a QFT on a lattice spacetime. Thus we have a family of physically equivalent bare theories, defined on lattices of decreasing spacing (as the cut-off limit is taken). For these distinct theories to capture the same physics, their couplings (and masses and field renormalization) must vary as the lattice spacing (or equivalently cut-off) varies. The RG

[1]The RG also considers the behaviour of the theory with respect to the variation of other mass quantities—in particular, with variation of the external momentum, in order to aid the calculation of high energy processes.

equations—the CS equation, the γ-function, and the β-functions—describe these variations.[2]

In arbitrary QFTs, every charge receives a β-function, and the RG equations become a set of coupled flow equations; even charges with zero magnitude in some bare theory might have non-zero β-functions, and so be 'turned on' as the lattice spacing changes. However, the basic RG idea is the same: by tuning the couplings as you go, one can find physically equivalent bare theories of the same form, but with different cut-offs.

The β-functions can be found (perturbatively) for specific theories, and hence CS equations can be solved.[3] For an 'initial' bare theory, with fixed cut-off and charges, $\bar{\Lambda}$ and $\bar{g}_0$, we find a solution of the form

$$\Gamma^{(N)}(p, g_0(\Lambda), \Lambda) = \xi(\bar{g}_0, \bar{\Lambda}/\Lambda) \cdot \Gamma^{(N)}(p, \bar{g}_0, \bar{\Lambda}) \qquad (13.10)$$

when $\bar{\Lambda}$ is scaled down to an equivalent theory with cut-off Λ by the RG. In particular we can trace the RG flow all the way to the arbitrary renormalization energy scale, μ:

$$\Gamma^{(N)}(p, g_0(\mu), \mu) = \xi(\bar{g}_0, \bar{\Lambda}/\mu) \cdot \Gamma^{(N)}(p, \bar{g}_0, \bar{\Lambda}). \qquad (13.11)$$

This expression gives us our interpretation of equation 13.7, linking bare and renormalized theories: identifying the bare theory of equation 13.7 with the initial theory of equation 13.11 (the RHS amplitude), implies that the renormalized theory of equation 13.7 should be identified with the LHS of equation 13.11. Thus the renormalized theory is the result of *applying the RG* to the bare theory, and field rescaling represents the rescaling effects of the RG: for a finite cut-off the renormalized theory contains the same physics as the bare theory, and is related to it by the RG. In fact we have only seen this idea in outline here, but it can be seen in more detail in specific cases (e.g., Le Bellac, 1991, sec. 7.3.2).

[2] A crucial point for anyone familiar with the RG from statistical physics is that $\Lambda \to \infty$ corresponds to a shrinking lattice spacing, whereas 'blocking' (e.g., Maris and Kadanoff, 1978) increases the lattice spacing, or integrates the cut-off down to *lower* values. That is, the RG in QFT and statistical physics typically concerns flows in opposite directions.

[3] The amplitude and field rescaling are calculable within perturbative renormalization, and so the CS equation can be imposed at each perturbative order to find an expansion for the β-functions. This method contrasts—but agrees with—Wilson's solid state approach in which the effect of integrating out high momentum modes is explicitly evaluated and absorbed into redefinitions of the parameters.

Let's now say a little more about the β-functions that describe the change in the charges as the cut-off varies. Consider a QFT with a number of (dimensionless) charges, $g_1, g_2, \ldots$, each with its own β-function, $\beta_1, \beta_2, \ldots$. The RG determines how each coupling must change as the cut-off varies in order to keep the same physics. Thus, we picture each theory as a point in a multidimensional 'parameter space' and the RG equations as generating a unique trajectory of equivalent theories, parameterized by the cut-off, through this space (e.g., Wilson and Kogut, 1974): see figure 13.1.[4]

The most interesting points of the coupling flows are those at which the β-functions vanish, the 'fixed points' (FPs) at which the couplings are constant with respect to changes in the cut-off. The most important example in QFT (in 4-dimensions) is the origin or 'Gaussian FP'—P in figure 13.1—describing free theories, with vanishing couplings. Locating a FP enables us to investigate how couplings vary with increasing Λ in its neighborhood, where their β-functions change smoothly from zero. There are two main cases to consider, $\beta(g)$ positive and $\beta(g)$ negative with increasing cut-off.

If $\beta(g)$ is positive then as the cut-off *increases* g will grow, leading the trajectory away from the FP, according to the RG transformation, and so one says that the FP is 'ultra violet (UV) unstable' for the parameter: as for g_1 in figure 13.1. Equivalently (in terms more normal to solid state physics) one can say that the coupling is 'irrelevant' at the FP, since as the cut-off *decreases*, the value of g_1 is driven to the FP value.

If, on the other hand, $\beta(g)$ is negative, as it is for g_2 in figure 13.1, then as the cut-off *grows* the charge will decrease towards the FP, which is termed 'UV stable' for the parameter. Again, in solid state terms such parameters are known as 'relevant', since for *decreasing* cut-off the the parameter can in principal take on any value, not that at the FP.

In fact, there is a further important classification of parameters. Relevant and irrelevant parameters vary as (negative and, respectively, positive) powers of the cut-off, but those parameters whose β-functions vanish to first order vary as logarithms of the cut-off; such couplings are called 'marginal'. Marginal couplings are far less sensitive to variations in the cut-off, but the basic picture is the same: irrelevant-like marginal couplings are driven away from the fixed point as the cut-off increases, but relevant-like marginals

[4]Note that RG dynamics are 'Aristotelian': the velocity—$dg/d(\ln \Lambda)$—is given directly by a field—$\beta(g)$—in (parameter) space.

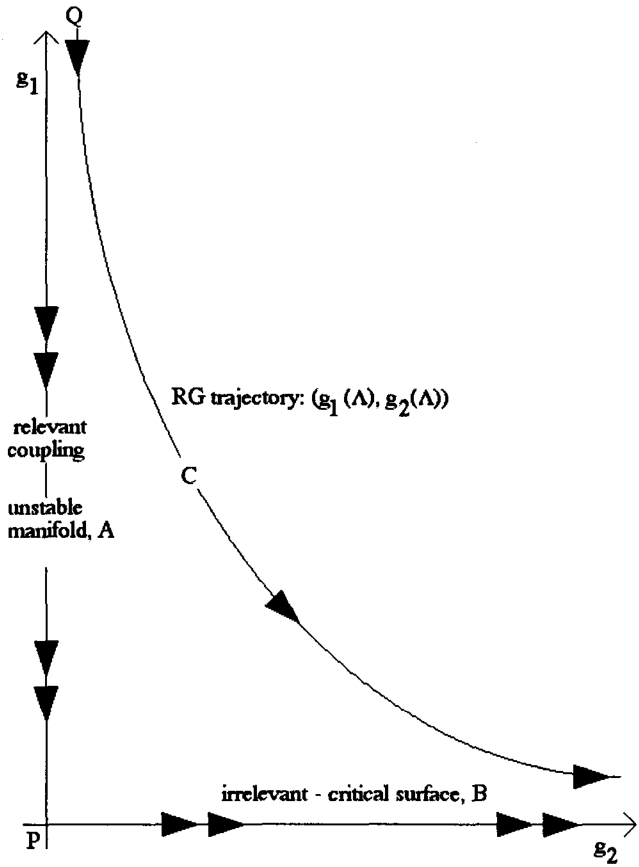


Fig. 13.1 *RG flows in parameter space for increasing* Λ

are driven towards it. Now, it turns out that marginal couplings (of the Gaussian FP) correspond to the 'renormalizable' couplings of perturbative renormalization, namely those with mass dimension $[\lambda] = 0$. Their physical significance makes it very interesting to study the higher order contributions to their β-functions to determine how exactly they behave under the RG. We will return to this question at the end of the section.

Assuming that the flows vary smoothly in parameter space, one can qualitatively picture the action of the RG on a theory which has both g_1 and

g_2 couplings and cut-off $\bar{\Lambda}$. As the cut-off varies, so do the two couplings, according to the RG equation, tracing out a curve of equivalent theories in parameter space. When $g_2 > g_1$ (in some appropriate sense) then the g_2 flow dominates—as in the first part of the curve drawn in figure 13.1—and vice versa. This trajectory is illustrated by C in figure 13.1.

Relative to any FP we can consider two disjoint submanifolds of parameter space: the 'critical surface', composed of all points corresponding to theories with *all* irrelevant(-like) parameters—B in figure 13.1—and the 'unstable manifold', or the the set of all theories with *only* relevant(-like) parameters—A. In other words, the critical surface consists of all theories that flow into the FP as the cut-off *decreases*, and the unstable manifold consists of all theories that lead into the FP as the cut-off *increases.*

The picture we have already gleaned from the RG allows us to start to address Teller's question: do divergences reveal a pathology in the theory itself, or in the approximation scheme? For consider the following (Weinberg, 1978): if a coupling does not have an UV stable FP, then any QFT containing it is unlikely to be well defined in the continuum limit. That is, as the cut-off for a bare theory containing such a parameter is taken to infinity, its flow line does not approach some limit—a FP with respect to which it is *relevant*—and some kind of pathology is likely. Following Weinberg (1978) one calls theories whose couplings have an UV FP 'asymptotically safe'; asymptotic freedom, which means that the Gaussian—free—FP is UV stable for the coupling, is thus a species of asymptotic safety. But note that the UV FP need not be the Gaussian FP: all that is necessary fir asymptotic safety is that the coupling be relevant with respect to some FP.[5]

Thus it becomes pressing to enquire whether our important physical theories such as φ^4, QED, QCD, and electro-weak theory, are asymptotically safe. Such theories have perturbatively renormalizable, which is to say marginal, couplings, and so the question is whether they have safe, relevant-like or unsafe, irrelevant-like behaviours. The perhaps surprising answer is that only QCD—and in general only non-Abelian gauge theories—is believed to be so. For other important QFTs we find the following:

> There are rigorous theorems suggesting that in four dimensions un-

[5]Note too that whether a coupling is relevant at a given FP or not depends on all the couplings in the theory. After all in general the RG flow is determined by a set of coupled flow equations.

less the coupling takes on the value zero, then φ^4 theory is not asymptotically safe (Fernandez, Fröhlch and Sokal, 1991, chap. 15, but see Weinberg, 1996, sec. 18.3, for dissent about their interpretation). (This need not pose a problem for the most important φ^4 theory, the Higg's field, because it is coupled to the electro-weak theory).

No UV FP of QED is known (the theory is IR stable at the Gaussian FP), and the consensus seems to be that there probably is none—it is most likely not asymptotically safe.

Electro-weak theory, despite being a non-Abelian gauge theory, has terms which seem to prevent asymptotic freedom, and so, if it has no other, UV, FP, won't be asymptotically safe (Zinn-Justin, 1993, sec. 33.4).

There are loopholes of various sizes for these theories, but the formal evidence is against their existing as continuum physics. These facts are not widely proclaimed by physicists, but in fact the possibility was first raised by Gell- Mann and Low in 1954.

13.4 Cartwright on Fundamentalism

Fundamental theories (of physics) have suffered greatly at the hands of Nancy Cartwright. First, in 'How the Laws of Physics Lie' (Cartwright, 1983—henceforth 'HLPL') she argued that the situations in which a regulated scientific account is possible do not literally fall under general laws: covering laws are simply false. Instead science uses descriptive *ceteris paribus* 'phenomenological' laws adapted to specific circumstances: laws that are not special cases of true fundamental laws, but which are thematically (and heuristically) organized by abstract general laws. Her arguments undermine the claim that general laws would have to be true in order to have the explanatory power that they do, and specifically aim to show that phenomenological laws are not logical consequences of fundamental laws. Instead, they are derived from causal models, which fundamental laws help us construct. Disunity in this context thus applies to the phenomenological laws; they are not just different realizations of some underlying truth, but autonomous.

In more recent work (in particular, 1994) Cartwright continues this attack, claiming that fundamental laws do not hold—even in her sense—across the whole of their expected domain. For instance, she cites examples of (non- relativistic) mechanical systems for which Newtonian mechanics offers no treatment. It is only in very specific 'contrived' situations that the fundamental laws hold, allowing a phenomenological treatment. But in many other situations we may know the causal powers involved—the action of gravitational forces for instance—but, she argues, this knowledge cannot be regimented within science in the way she suggested earlier: we can produce no model from which to derive a phenomenological law. If one accepts her earlier arguments then this conclusion is perhaps not such a surprise, for if general laws are merely tools for organizing and constructing phenomenological treatments, then one might well expect their instrumental value to vary from case to case, and to give out altogether eventually. In this new context the disunity is even greater: it isn't even the case that all phenomena within some domain can be modeled by a covering theory, true or not. Literally, nature cannot be united by science, even instrumentally.

Now, there are a number of things to be said against Cartwright's program, but doing so is not my goal here. Instead I want first to propose that perturbative renormalization raises difficulties for a realist claim that QFT is a true fundamental theory, and that these can best be articulated in terms of her arguments. Then I want to use our work on the RG to try to dispell those problems. To set up the arguments I will devote the remainder of this section to a discussion of a couple of ideas from HLPL in more detail: in particular the difference between fundamental and phenomenological laws, and what purported relationship between them is criticised by Cartwright.

One common use of the expression 'fundamental law' is to refer to the few best theories of the day: in our case the standard models of high energy and cosmological physics (or perhaps their unification within string theory). Cartwright sometimes has this meaning in mind, but the examples that she uses show that the idea is more general. For instance, she discusses in HLPL—as fundamental—fluid dynamics, Newtonian mechanics, Newtonian gravity, Coulomb's Law, a law for midband gain of amplifiers and a non-relativistic derivation of the Lamb shift. So, her complaint is not just with 'ultimate laws' but with any laws putatively covering some suitably broad domain.

Her attacks on such laws are of a piece with other contemporary ideas in philosophy of science. For instance, Ian Hacking (1982) argues that there is no single 'theory of the electron' used in the construction of a polarized electron gun. Instead there are different pieces of engineering 'lore' that experimentalists use to build the machine. Clearly this claim is tantamount to saying that any fundamental theory of electrons is a lie, though perhaps useful for organizing an array of causal knowledge.

In this case the notion of 'fundamental' is not intended to imply complete universality or even strict truth within a limited domain. Newtonian mechanics is fundamental in Cartwright's sense, but it requires no philosophical argument to establish that it is not universally true, rather than approximately true in a suitable domain. Thus Cartwright must be attacking the view that fundamental theories are even approximately true in their putative domains. What is ingenious about her arguments is that they do not invoke traditional worries about the cogency of 'approximate truth', but rather question the very idea that any kind of truth (even 'empirical adequacy' in van Fraassen's, 1976, sense) has any business consorting with fundamental theories at all.

An important part of HLPL is concerned with disputing the claim that explanations require truth, and setting up an alternative model of explanation. But at the heart of the book (Essay 6) is an argument against a realist account of the relation between fundamental and phenomenological laws—the generic- specific (GS) account:

> "... when fundamental laws explain a phenomenological law, the phenomenological law is deduced from the more fundamental in conjunction with a description of the circumstances in which the phenomenological law obtains. The deduction shows just what claims the fundamental laws make in the circumstances described." (Cartwright, 1983, 103)

The GS model provides the realist with two things. First, there is a metaphysically tidy account of what is 'fundamental' about fundamental laws. Putting the story in terms of deduction means that (with Grünbaum, 1954, 14) we don't have to think of the laws of physics as *causing* the phenomenological laws. The idea is that 'phenomenological laws are what the fundamental laws amount to in the circumstances at hand' (HLPL 103); a fundamental law covers every situation within its domain, and all that is left to do is to find out what they dictate on specific occasions. The second thing that the GS account provides is a model of indirect confirmation; if many empirically confirmed phenomenological laws are specific consequences of some general law, then the (best) explanation of the empirical truth of the specific laws is that the fundamental law is true. Cartwright attacks both points by disputing the claim that phenomenological laws are logical consequences of our fundamental laws. Of course, we offer 'derivations' of specific laws from general ones, but in her model of scientific explanation, there is no requirement that these derivations be deductions. In particular, the approximations used in such derivations are held by her not to represent sound inferences, but to serve to ensure that we get true consequences from our false laws.

With these basic ideas of Cartwright's philosophy in mind I now want to describe the problem that renormalization poses for fundamentalism about QFT. In the next section I will first characterize perturbative renormalized QFT as a kind of phenomenological law compared to a fundamental, exact QFT. Second, I will explain how the need for renormalization is a problem for the realist claim that the GS account captures the relationship between the two theories, and how this problem is analogous to those raised by Cartwright. In the final section I will use the RG analysis to defend the GS model in this context.

13.5 Renormalization as Grounds for Disunity

A reasonable way to fit the fundamental-phenomenological distinction to high energy physics is to view exact path integral QFT as the fundamental theory and particular scattering cross-sections calculated from it using the techniques of perturbation theory and renormalization as phenomenological laws. To paraphrase, the idea is that the amplitudes are what an exact QFT amounts to in the scattering process at hand. Let me just briefly comment on this picture.

First, there are two important senses in which one might say that path integral QFT is not 'fundamental'. One can identify such a theory by writing down an appropriate Lagrangian for the field, but in an important way the field so defined is heuristic. Namely, renormalization aside, such a system, with point field operators, is not well-defined on a continuum; if the theory exists in a rigorous sense then operators must be 'smeared' using test functions. The fundamental theory is supposed to mean the exact amplitudes extracted from the generating functional, but it is not clear that this notion makes literal sense. We shall just ignore this potentially serious point.[6]

I have already hinted at a second sense in which QFTs may not be 'fundamental': namely they may just be low energy approximations to some totally different 'final theory', such as string theory. Again, given Cartwright's understanding of 'fundamental' law, this is no particular problem, for even if string theory ultimately is the correct account of the physical realm, it will always be the case that the QFTs of our standard model are putatively 'true in their domain', as Newtonian mechanics is in its domain. One might also question the sense in which scattering amplitudes are 'phenomenological'. After all, in the philosopher's sense such things are not 'directly observable', but only indirectly, through the use of detectors and computer analysis. But this is unproblematic in Cartwright's analysis, for 'phenomenological' is to be used in the broad physicist's sense, and not as a synonym for 'visible to the naked eye'.

The realist picture of the situation is thus that the fundamental QFT holds in the appropriate domain of high energy physics, and that, GS-style, the phenomenological scattering amplitudes are (approximations to) deductive consequences of it. The problem is that renormalization obscures the logical relationship between fundamental and phenomenological QFT. To see this, let's first briefly review Cartwright's explicit arguments.

As it happens, Cartwright (HLPL Essay 6) does discuss a renormalized theory— non-relativistic QED—but she is not concerned with renormalization itself, but rather with the approximations that are applied to the theory. The practicalities of concrete experiments are such that neither exact descriptions of the circumstances nor strict deductions are possible. One must generalize and idealize the situation, and one must 'guess' how a derivation would go, jumping over logical steps with inspired leaps that one hopes have much the same consequences as a strict deduction; this

[6] As a result, one may question on how rigorous a footing the later considerations of this paper can be placed. Moreover, as Andrew Wayne forcefully argued in Bielefeld, the reasoning described here only seems to show the existence of continuum QFTs in the limit of a certain sequence of lattice theories. And one might ask, as he did, whether that really demonstrates the existence of continuum QFTs that can 'stand on their own two feet'. It seems that the only completely satisfying response to such worries would be to give a proof within axiomatic QFT of a suitable existence theorem. Not only would this project far exceed the ambitions of this paper (or this author), physicists, as far as I can tell, have wildly differing opinions about the possibility of proving such theorems, on both logical and practical grounds. The considerations offered here then constitute the most rigorous, most fruitful way of pursuing our questions.

business of finding suitable approximations is of course absolutely central to the successes of science. For the GS model to hold, it must be the case that our inspired leaps genuinely take us in the same direction as a strict calculation. We have to have grounds for believing that the logical steps are 'in place' behind the scenes of our derivations, if we want to believe the GS view.

Cartwright offers two kinds of example to undermine the claim that the logical steps are in the background at all. First, once we make approximations, we can end up with results that are more accurate than those obtained by strict derivations. Second, in general, there are a number of different possible approximations, and which one is most accurate is decided, not bottom-up by the fundamental theory, but only top-down by the phenomena we wish to capture. Now, as Cartwright concedes, it is not clear that these points prove that phenomenological laws, derived approximately, are not (approximate) logical consequences of fundamental laws. Instead, Cartwright sees the argument as showing that a burden of proof rests on the proponent of the GS model: 'Realists can indeed put a gloss on these examples... But the reason for believing the gloss is ... realist metaphysics...' (HLPL 126-7). In other words, if she is right, the typical derivation of a phenomenological law does not provide grounds for thinking it is a deductive consequence of the fundamental law.

I won't analyze her arguments here, but rather wish to point out that phenomenological QFT raises, in spades, exactly the same kind of difficulties for the GS account. The problem is not the use of approximations, for in QFT the main approximation is a perturbative expansion in powers of a small charge, which is *prima facie* well justified. Rather, the need to renormalize in perturbation theory raises far graver doubts that phenomenological cross- sections are deductive consequences of exact path integral QFT than any of the examples she produces. *Prima facie*, the case is simply this: phenomenological results differ from anything derived in an (approximate) deductive manner by an infinite amount, until one renormalizes. Now, we have seen that one uses a cut-off during renormalization, so all the mathematics involved is formally legitimate. And it is certainly not the case that we can get any results whatsoever by renormalizing: for a renormalizable theory, we need only fix a finite number of parameters to calculate an infinite range of predictions, so a renormalized QFT has genuine predictive power. But it is hard to see from the technique how throwing away infinities, even in a formally legitimate manner, can count as a deduction. Note that in this case the argument is apparently stronger than Cartwright's. We haven't just neutralised positive arguments for the GS model—the need for perturbative renormalization provides specific grounds for doubting it. In the final, following section I will explain how our RG analysis counters these grounds, making the GS model tenable in QFT.

13.6 Renormalized QFT as a Fundamental Theory

Let's put the issue this way: we can see that perturbation theory is (in principle) justified as a method for obtaining approximations to exact QFT. But we also know that it is not sufficient alone, and naive application must be supplemented with renormalization, which raises the question of what possible role renormalization could play in our (approximate) deduction. Wilson (Wilson and Kogut, 1974, sec. 12) showed how to use the RG to provide an answer. As before, we work in a multi-dimensional parameter space, and once again picture two dimensions only. And let us assume again that these two dimensions are parameterized by a relevant(-like) and an irrelevant(-like) coupling. But let us this time consider the RG flow as the cut-off *decreases not increases.* In this case, the flows of figure 13.1 are reversed: relevant parameters flow away from the FP and irrelevant parameters towards it (this is the usual statistical physics picture).

Suppose then that we are interested in a QFT with a particular set of (non-zero) couplings with some given values—a point in parameter space, P in figure 13.2. To define some physics we will also need to specify a cut-off (or equivalently a lattice spacing) for this theory; say, $\bar{\Lambda}_0$. Now, there is no reason to suppose that the couplings at P will define a finite QFT for arbitrarily large values of the cut-off. In the limit that $\bar{\Lambda}_0 \rightarrow \infty$ modes of arbitrarily large momenta will contribute to the theory, in general leading to divergences. Such divergences make themselves known in perturbative QFT, but note that the point is equally applicable to exact QFT: we should expect that the amplitudes of an exact QFT, defined by its path integrals, are not well-defined if it has no cut-off.

A generic point such as P may not define a finite theory in the infinite cut-off limit, but a point—say Q of figure 13.2—on the critical surface--curve B—does. In broad strokes, we can see that criticality is required by reasoning as follows. First, it turns out that the propagator of a QFT with

Hamiltonian H and cut-off $\bar{\Lambda}$ is an exponential function for large distances:

$$\int_0^{\bar{\Lambda}} \mathcal{D}\varphi\ \varphi(0)\varphi(x)e^{-H/\hbar} \sim e^{-x/\xi}, \tag{13.12}$$

where ξ is a constant associated with the theory, the 'correlation length' (in dimensionless units). The critical surface, B, then can be given an equivalent definition to that already given (as the set of theories with only irrelevant couplings) as the set of points for which $\xi = \infty$.[7] Now the propagator, loosely speaking, is the quantum probability for a particle at 0 to be found at x, and so the distance ξ, or rather $\xi/\bar{\Lambda}$ in physical distance units, is a measure of the uncertainty in a particle's location. If we take the physical—renormalized— mass, μ_r, as a measure of the momentum uncertainty of a particle, then the uncertainty relations give $\mu_r \cdot \xi/\bar{\Lambda} = 1$ or $\mu_r = \bar{\Lambda}/\xi$. Thus we find a finite physical mass with an infinite cutoff only if the theory is critical. Put another way, continuum QFTs correspond the classical statistical systems undergoing second order phase transitions.

So suppose we've found that the interactions of some quantum field are described by some set of couplings. For typical values of these couplings—those we measure for example!—we find that there is no QFT if there is *no cut-off: equivalently, there is no* QFT *with those values of the couplings* unless space is a lattice, so there is no continuum QFT with those values. However, we can use what we know about parameter space to define a theory that has the desired non- zero couplings, but with different values from those we originally assigned. The trick is to consider a family of theories, all with the desired couplings, parameterized by the cut-off $\bar{\Lambda}$, such that in the limit $\bar{\Lambda} \to \infty$ the family tends to criticality. In parameter space as shown in figure 13.2 this family is represented by the curve A, parameterized by $\bar{\Lambda}$, with point Q on the critical surface as its limit. That is, we make the desired couplings functions of the cut-off, constrained to take on critical values in the continuum limit. P is now an essentially arbitrary point on this curve, with essentially arbitrary cut-off $\bar{\Lambda}_0$.

[7] In solid state physics the propagator corresponds to the correlation function, the measure of correlation between the parts of the system a distance x apart. Then the critical surface describes systems with infinite range correlations, which are interpreted as 'critical' systems undergoing second order phase transitions. And then the fact that all points on the surface are driven to the FP by the RG—which of course preserves their physical consequences—explains the 'universality' of critical phenomena in widely varying systems: the basic physics of criticality is determined by the FP for any kind of theory located on the critical surface.

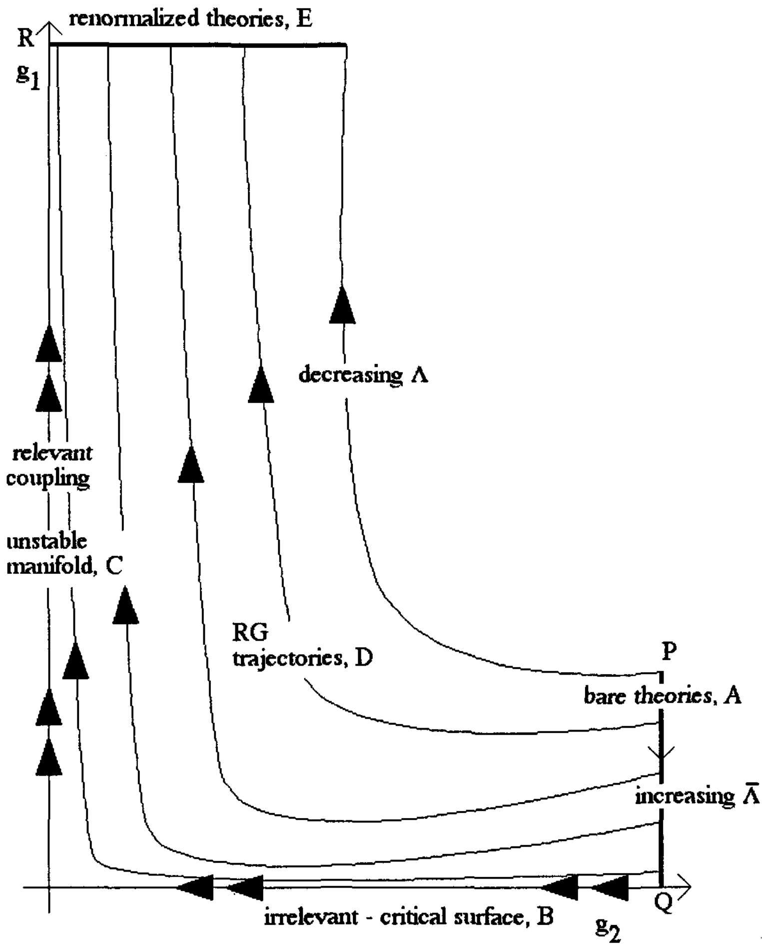


Fig. 13.2 Renormalizing a sequence of bare theories

Let's be clear, this 'trajectory' is *not* one generated by the RG: our prescription for the family involved no demand that its members be physically equivalent, as our construction in section 13.3 did. Instead, these theories correspond to the sequence of *bare theories* in perturbative renormalization as the cut-off is taken to infinity—the Γ_0 of equations 13.5 and 13.6 as $\bar{\Lambda} \to \infty$. And because this family tends to the critical surface, where we know continuum physics is possible, we expect that every point on the curve, with its given couplings and cut-off, corresponds to a finite bare theory, even in the infinite limit. (Why bother with the family rather than just picking a theory with infinite cut-off on the critical surface? Because extracting physical consequences requires canceling infinities, and this can only make sense as a limiting process, as we'll see.)

Moreover, we can give a graphical, heuristic proof that this sequence of theories does indeed yield well-defined physics in the continuum limit (we still follow closely the argument given by Wilson and Kogut, 1974, sec. 12, but see Polchinski, 1984, for a rigorous argument along these lines, at least for a simple QFT). First, note that for a given value of $\bar{\Lambda}$ we have a cut-off QFT of our family, to which we can apply the RG, producing a sequence of physically equivalent theories, parameterized by their cut-offs Λ. That is, solving equations 13.8 and 13.9 for each of the bare theories generates a series of trajectories, as in equations 13.7 and 13.10. Graphically, applying the RG to each point on A we induce the family of curves, D in figure 13.2. Each curve is labelled by its 'initial' cut-off $\bar{\Lambda}$ while the points on any particular curve are labelled by Λ, or if we make the change of variable $\Lambda \equiv e^{-t}\bar{\Lambda}$, by t. And again, all the points on a single curve describe cut-off theories with the same physics, though distinct curves do not (since their intial points correspond to different physics).

And of course these RG trajectories will follow the paths indicated. They cannot cross, for a trajectory is determined (by equations 13.8 and 13.9) completely by a single point in parameter space (their 'velocities' in parameter space are not degrees of freedom). And as they start ever closer to the critical surface (B) their initial behaviour is dominated by the variation of the irrelevant parameter, and as they get closer to the unstable manifold (C), the flow is dominated by the variation in the relevant parameter (as we discussed in connection with figure 13.1).

Now define some fixed renormalization scale μ, and consider for each RG trajectory just defined—so for each $\bar{\Lambda}$—the unique theory obtained by rescaling $\bar{\Lambda}$ to $\mu = e^{-t}\bar{\Lambda}$ using the RG. This set of 'renormalized' theories form the curve E in figure 13.2. Each renormalized theory is equivalent to its corresponding bare theory, though it has a different (and lower by a factor of e^{-t}) cut-off. Each renormalized theory has the same cut-off μ as any other, though none are physically equivalent, since their corresponding bare theories are not. To reiterate, we have used the RG and a renormalization scale to generate a family of renormalized theories, E in figure 13.2, from our family of bare theories, A.

Finally we have the pieces in place to confirm that the family of bare theories describes well-defined physics in the continuum limit. The physics of every bare theory is captured by the physics of the corresponding renormalized theory, which by construction is rendered finite by a fixed cut-off, hence the physics of the limit of the bare theories, point Q, is captured by the limit of the renormalized theories, point R, which is itself finite. And that, essentially, is that.

Of course, one might question whether the assertion that the limit of the renormalized theories exists is justified: indeed, justifying this assertion is the main step of the proof. One might in particular have the following worry. As the bare theory parameter $\bar{\Lambda}$ grows, the more work, the more rescaling, must be done by the RG to reduce the cut-off to our chosen value μ. Thinking of t as a 'temporal' metric, since the RG must be applied so that $\Lambda = e^{-t}\bar{\Lambda} = \mu$, the greater the initial cut-off, the greater t is—the longer it takes to reach the renormalized theory. And in the limit it takes an infinite amount of time. In this case, what justifies our assumption that the RG trajectory has an *end-point* in the $\bar{\Lambda} \to \infty$ limit? That there really is a finite renormalized theory in that limit?

It is the topology of parameter space that justifies this assumption. As the limit is taken, the trajectory starts ever closer to the critical surface, B, and so the variation of the relevant parameter with t dominates ever more, and so the resulting trajectory follows the critical surface ever more closely to the FP, before following the unstable manifold, C, away. But at a FP, the velocity of a RG trajectory is by definition zero: the trajectory does not move in parameter space at all as t varies. So as the limit is taken (by continuity) the ever increasing amounts of time it takes to rescale the cut-off down to the renormalization scale are 'used up' in the ever increasing amounts of time it takes the trajectory to get past the FP. In the limit, the infinite amount of time it takes to rescale an infinite cut-off down to a finite value is spent actually at the FP, and 'then', as it were, the trajectory passes down the unstable manifold until the cut-off takes on the renormalized value. And so the crucial step in the argument that we have well-defined continuum physics in the limit is justified.

In summary, we know that the $\bar{\Lambda} \to \infty$ limit of bare theories, on the critical surface, describes good physics, because it describes the same physics as a manifestly finite theory on the unstable manifold.

It should be apparent that not only have we seen a graphical proof-sketch of the existence of continuum QFTs, but we have also just seen a graphical portrayal of renormalization as described earlier for perturbation theory. We pick a theory by the values of its couplings, here represented by P. We find that with no cut-off it is pathological, so we give it a cut-off, $\bar{\Lambda} = \bar{\Lambda}_0$ as in equation 13.3. Then we redefine the parameters to make them $\bar{\Lambda}$ dependent, which gives us a family of bare theories as in equation 13.4. Each of these bare theories generates a trajectory of equivalent theories according to the RG, and in particular one with a given renormalized cut-off μ, as in equation 13.5. Finally we take the continuum limit of the bare theories (the RHS of equation 13.6) to obtain the finite physics of the corresponding renormalized theory, (the LHS of equation 13.6).

In perturbative renormalization we of course make the couplings functions of the cut-off, subject to the constraint that they cancel out divergences. How does this constraint correspond to the demand that the limit of the bare theories be a point on the critical surface? The story requires that any finite renormalized continuum theory live on the unstable manifold—so that the limiting RG trajectory passes through the FP. And the demand that the renormalized theory be on the unstable manifold means that the bare theory must lie on the critical surface—the topology is such that only trajectories starting on the critical surface end up on the unstable manifold.[8]

Thus we have seen that there can be continuum QFTs *and* what role perturbative renormalization plays in finding them. Arbitrary QFTs are not well-defined in the continuum, only those on the critical surface. Renormalization provides a way to explore parameter space for such a theory and delivers its physics: by demanding that the couplings cancel divergences we demand a critical theory—picking a renormalization scale determines the renormalized theory that captures the physics. That is, renormalization plays a crucial, interpretable role in perturbation theory, as a necessary step in the derivation of QFT phenomena from fundamental theory. It poses no threat to the GS account of QFT.

And thus we also see the answer to Teller's question. Divergences in QFT do not arise from the perturbation scheme, but from the fact that QFTs only have continuum formulations for very specific values of their couplings. As such, assuming that a theory is aymptotically safe, the prob-

[8]Remember that we are talking about a trajectory defined as the limit of a sequence of trajectories that do not intersect the critical surface.

lem is not terminal but can be fixed using renormalization to tune it to a suitable point. If, however a theory is not asymptotically safe, then for no values of its parameters will it live on the unstable manifold of a FP, and renormalization as described here will not be possible.

Acknowledgements

This material is based upon work supported by a faculty fellowship from the Center of the Humanities at the University of Illinois at Chicago and by the National Science Foundation under grant number SES-0004375.

References

Le Bellac, M. (1991): *Quantum and Statistical Field Theory.* Oxford University Press, Oxford.

Cartwright, N. (1983): *How the Laws of Physics Lie.* Oxford University Press, Oxford.

Cartwright, N. (1994): Fundamentalism vs the Patchwork of Laws. *Proceedings of the Aristotelian Society* **93**, 279-292.

Cheng, T.-P. and Li, L.-F. (1984): *Gauge Theory of Elementary Particle Physics.* Oxford University Press, Oxford.

Fernandez, R., Frolich, J. and Sokal, A. D. (1992): *Random walks, critical phenomena, and triviality in quantum field theory.* Springer-Verlag, Berlin.

van Fraassen, B. C. (1996): *The Scientific Image.* Oxford University Press, Oxford.

Gell-Mann, M. and Low, F. E. (1954): Quantum Electrodynamics at Small Distances. *Physical Review* **95**, 1300.

Grünbaum, A. (1954): Science and Ideology. *The Scientific Monthly* 14.

Hacking, I. (1982): Experimentation and Scientific Realism. *Philosophical-Topics* **13**, 71-88.

Huggett, N. and Weingard, R. (1996): The Renormalization Group and the Effective Field Theory Programme. *Synthese* **102**, 171-194.

Maris, H. J. and Kadanoff, L. P. (1978): Teaching the Renormalization Group. *American Journal of Physics* **46**, 652-657.

Polchinski, J. (1984): Renormalization and Effective Lagrangians. *Nuclear Physics* **B231**, 269-295.

Ryder, L. H. (1985): *Quantum Field Theory.* Cambridge University Press, Cambridge, UK.

Shenker, S. (1984): Field Theories and Phase Transitions. In J. B. Zuber and R. Stora, eds., *Recent Advances in Field Theory and Statistical Mechanics.* Elsevier Science, New York.

Teller, P. (1995): *An Interpretive Introduction to Quantum Field Theory.* Princeton University Press, Princeton.

Weinberg, S. (1978): Critical Phenomena for Field Theorists. In A. Zichichi, ed., *Understanding the Fundamental Constituents of Matter.* Plenum Press, New York. 1-52.

Wilson, K. G. and Kogut, J. (1974): The Renormalization Group and the ϵ-Expansion. *Physics Reports* **12**, 75-200.

Zinn-Justin, J. (1993): *Quantum Field Theory and Critical Phenomena.* Oxford University Press, Oxford.

This page is intentionally left blank

PART 4

Gauge Symmetries and the Vacuum

This page is intentionally left blank

Chapter 14

The Interpretation of Gauge Symmetry

Michael Redhead
LSE, London

Abstract. *In its most general sense gauge freedom in the mathematical description of a physical system refers to ambiguity in that description. The general relationship between ambiguity of representation and physical symmetries is explained. The case of surplus structure is examined, where there are more degrees of freedom in the mathematical description than in the physical system itself. This leads to the concept of a constrained system in which the equations of motion contain arbitrary functions representing the gauge freedom. As a result the time-evolution of the mathematical degrees of freedom is indeterministic. Examples of constrained Hamiltonian systems are provided by the free-field Maxwell equations for the electromagnetic field, and the equations of canonical general relativity. In a still more restricted sense gauge freedom refers to the situation in so-called Yang-Mills gauge theories of elementary particle interactions, where the form of possible interactions is constrained by a principle of local gauge symmetry referring to the generalised phases associated with the wave functions of the matter fields. The interpretation of Yang-Mills symmetry involves a trilemma between the indeterminism associated with a realistic interpretation of the gauge potentials, the nonlocality associated with attempts to formulate the theory in terms of purely gauge-invariant qualities and a potentially mysterious Platonist-Pythagorean role for purely mathematical constructions in controlling the physical world on an antirealist construal of the potentials. More recent developments involving BRST symmetry are discussed in this context.*

14.1 Introduction

The term "gauge" refers in its most general everyday connotation to a system of measuring physical quantities, for example by comparing a physical magnitude with a standard or "unit". Changing the gauge would then refer to changing the standard. The original idea of a gauge as introduced by Weyl in his (1918) in an attempt to provide a geometrical interpretation of the electromagnetic field was to consider the possibility of changing the standard of "length" in a four-dimensional generalization of Riemannian geometry in an arbitrary local manner, so that the invariants of the new geometry were specified not just by general coordinate transformations but also by symmetry under conformal rescaling of the metric. The result was, in general, a nonintegrability or path dependence of the notion of length which could be identified with the presence of an electromagnetic field. In relativistic terms this meant that unacceptably, the frequencies of spectral lines would depend on the path of an atom through an electromagnetic field, as was pointed out by Einstein.

With the development of wave mechanics the notion of gauge invariance was revived by Weyl himself (1929) following earlier suggestions by Fock and by London, so as to apply to the nonintegrability of the *phase* of the Schrödinger wave function, effectively replacing a scale transformation $e^{\alpha(x)}$ by a phase transformation $e^{i\alpha(x)}$. Invariance under these local phase transformations, as contrasted with constant global phase transformations, necessitated the introduction of an interaction field which could be identified with the electromagnetic potential, a point of view which was particularly stressed by Pauli (1941). The extension of this idea to other sorts of interaction was introduced by Yang and Mills in their (1954) (although mention should be made of the independent work of Shaw (1954) and the proposals made in an unpublished lecture by Oskar Klein in 1938). The extension to a gauge theory of gravitation was considered by Utiyama (1956). The great advantage of gauge theories was that they offered the possibility of renormalizability, but this was offset by the fact that the interactions described by gauge fields were carried by massless quanta and so seemed inappropriate to the case of the short-range weak and strong interactions of nuclear physics. In the case of the weak interactions this defect was remedied by noticing that renormalizability survived the process of spontaneous symmetry breaking that would generate effective mass for the gauge quanta, while the key to understanding strong interactions as a gauge theory lay in the development of the idea of "asymptotic freedom", expressing roughly the idea that strong interactions were actually weak at very short distances, effectively increasing rather than decreasing with distance.

With this brief historical introduction we turn to consider the fundamental conceptual issues involved in gauge freedom and the closely associated idea of gauge symmetry.

14.2 The Ambiguity of Mathematical Representation

As we have seen the term gauge refers in a primitive sense to the measurement of physical magnitudes, i.e. of associating physical magnitudes with mathematical entities such as numbers. Of course the numerical measure is not unique, varying indeed inversely with the magnitude of the unit chosen. Both the unit and the measure can, with some confusion, be referred to as the gauge of the quantity, in everyday parlance.

We now want to generalize this usage by referring to the mathematical representation of any physical structure as a gauge for that structure. By narrowing down this very general definition we shall focus in on more standard definitions of gauge in theoretical physics, such as the gauge freedom of constrained Hamiltonian systems and Yang-Mills gauge symmetries.

But let us start with the most general concept.[1] Consider a physical structure P consisting of a set of physical entities and their relations, and a mathematical structure M consisting of a set of mathematical entities and their relations, which represents P in the sense that M and P share the same abstract structure, i.e. there exists a one-one structure-preserving map between P and M, what mathematicians call an isomorphism. In the old-fashioned statement view of theories, P and M could be regarded as models for an uninterpreted calculus C, as illustrated in figure 14.1. On the more modern semantic view theories are of course identified directly with a collection of models such as P. We do not need to take sides in this debate. For our purposes we need merely to note that P does not refer directly to the world, but typically to a "stripped-down", emasculated, idealized version of the world. (Only in the case of a genuine Theory of

[1]The following account leans heavily on Redhead (2001).

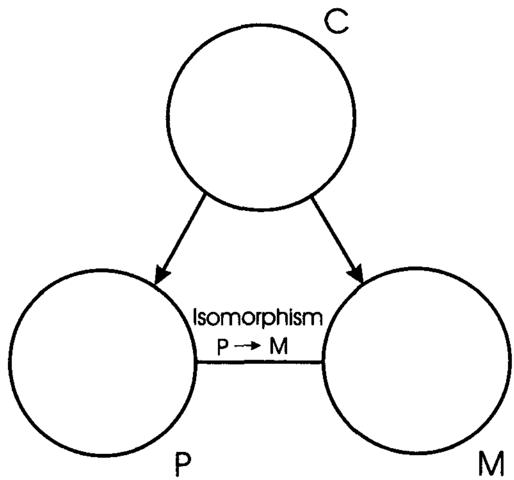


Fig. 14.1 *A physical structure P and a mathematical structure M are isomorphic models of an uninterpreted calculus C.*

Everything would there be a proposed isomorphism between the *world* and a mathematical structure.)

In our new terminology we shall call M a gauge for P (another way of expressing the relationship between P and M, would be to say that M "coordinatizes" P in a general sense).

In general there will be many different gauges for P. Consider, as a very elementary example, the ordinal scale provided by Moh's scale of hardness. Minerals are arranged in order of 'scratchability' on a scale of 1 to 10, i.e. the physical structure involved in ordering the hardness of minerals is mapped isomorphically onto the finite segment of the arithmetical ordinals running from 1 to 10. But of course we might just as well have used the ordinals from 2 to 11 or 21 to 30 or whatever. The general situation is sketched in figure 14.2, which shows two maps x and y which are isomorphisms between P and distinct mathematical structures M_1 and M_2. Of course M_1 and M_2 are also isomorphically related via the map $y \circ x^{-1} : M_1 \to M_2$ and its inverse $x \circ y^{-1} : M_2 \to M_1$.

But how can the conventional choice between M_1 and M_2 as gauges for P have any *physical* significance? To begin to answer this question we introduce the notion of a symmetry of P and its connection with the gauge

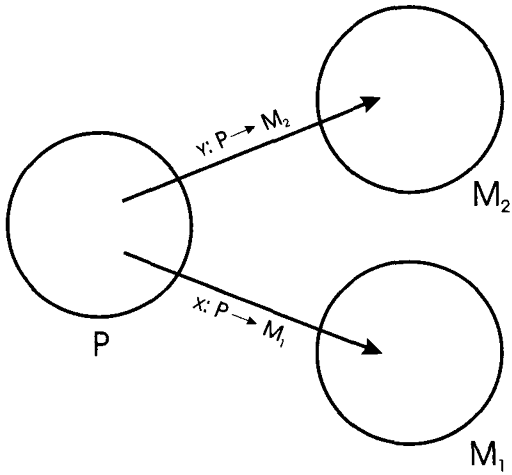


Fig. 14.2 *Ambiguity of gauge. M_1 and M_2 are distinct mathematical structures each of which represents P via isomorphisms x and y respectively.*

freedom in the generalized sense we have been discussing.

14.3 Symmetry

Consider now the case where the ambiguity of representation (the gauge freedom) arises within a *single* mathematical structure M. Thus we consider two distinct isomorphisms $x : P \to M$ and $y : P \to M$, as illustrated in figure 14.3.

Clearly the composite map $y^{-1} \circ x : P \to P$ is an automorphism of P. This is referred to by a mathematician as a point transformation of P and by physicists as an *active* symmetry of P. The composite map $y \circ x^{-1} : M \to M$ is a "coordinate" transformation or what physicists call a *passive* symmetry of P. It is easy to show that *every* automorphism of P or M can be factorized in terms of pairs of isomorphic maps between P and M in the way described. It is, of course, not at all surprising that the automorphisms of P and M are themselves in one-one correspondence. After all, since P and M are isomorphically related, they share the same abstract structure, so the structural properties of P represented by the

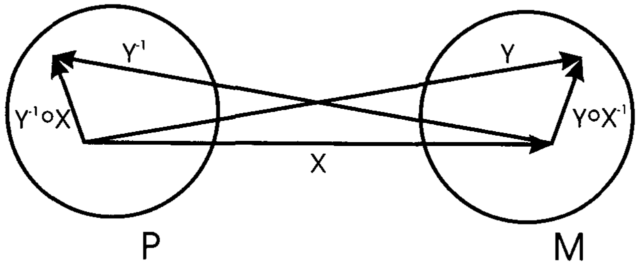


Fig. 14.3 *x and y are two distinct isomorphisms between P and M. Then $y^{-1} \circ x : P \to P$ is an automorphism of P and $y \circ x^{-1} : M \to M$ is an associated automorphism of M.*

symmetries of P can be simply read off from the corresponding symmetries of M.

Now the symmetries of P express very important structural properties of P, and we can see how they are related to the gauge freedom in this very important special case where the ambiguity of representation is within a *single* mathematical structure M.

The gauge freedom represented in figure 14.2 does not, in general, have physical repercussions related to symmetry. For example, in the case of Moh's scale of hardness, there simply are no non-trivial automorphisms of a finite ordinal scale.

We now want to extend our discussion to a more general situation, which frequently arises in theoretical physics and which we introduce via a notion we call "surplus structure".

14.4 Surplus Structure

We consider now the situation where the physical structure P is *embedded* in a larger structure M' by means of an isomorphic map between P and a substructure M of M'. This case is illustrated in figure 14.4.

The relative complement of M in M' comprises elements of what we shall call the surplus structure in the representation of P by means of M'. Considered as a structure rather than just as a set of elements, the surplus structure involves both relations among the surplus elements and relations between these elements and elements of M.

A simple example of this surplus structure would arise in the familiar use

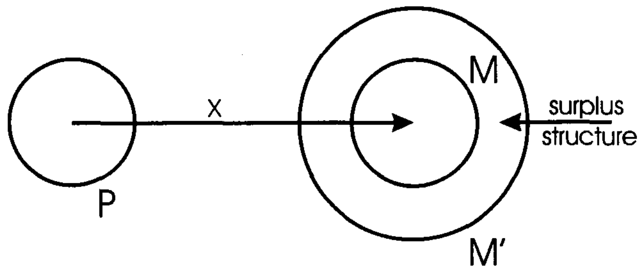


Fig. 14.4 *$x : P \rightarrow M$ is an embedding of P in the larger structure M'.*

of complex currents and impedances in alternating current theory, where the physical quantities are embedded in the wider mathematical structure of complex numbers. Another example is the so-called S-matrix theory of the elementary particles that was popular in the 1960s, in which scattering amplitudes considered as functions of real-valued energy and momentum transfer were continued analytically into the complex plane and axioms introduced concerning the location of singularities of these functions in the complex plane were used to set up systems of equations controlling the behaviour of scattering amplitudes considered as functions of the real physical variables. This is an extreme example of the role of surplus structure in formulating a physical theory, where there was no question of identifying any physical correlate with the surplus structure.

In other examples the situation is not so clear. What starts as surplus structure may come to be seen as invested with physical reality. A striking example is the case of energy in 19th century physics. The sum of kinetic and potential energy was originally introduced into mechanics as an auxiliary, purely mathematical entity, arising as a first integral of the Newtonian equations of motion for systems subject to conservative forces. But as a result of the formulation of the general principle of the conservation of energy and its incorporation in the science of thermodynamics (the First Law) it came to be regarded as possessing ontological significance in its own right. So the sharp boundary between M and the surplus structure as illustrated in figure 14.4 may become blurred, with entities in the surplus structure moving over time into M. Another example would be Dirac's hole theory of the positron, allowing a physical interpretation for the negative-energy solutions of the Dirac equation.

Ambiguities in representation, i.e. gauge freedom, can now arise via automorphisms of M' that reduce to the identity on M, i.e. the transformations of representation act non-trivially only on the surplus structure. Nevertheless such transformations can have repercussions in controlling the substructure M and hence the physical structure P. This is the situation that arises in Yang-Mills theories which we shall describe in section 14.6. But first we shall make a short digression to discuss the example of constrained Hamiltonian systems, of which free-field electromagnetism is a very important special case.

14.5 Constrained Hamiltonian Systems[2]

The idea of surplus structure describes a situation in which the number of degrees of freedom used in the mathematical representation of a physical system exceeds the number of degrees of freedom associated with the physical system itself. A familiar example is the case of a constrained Hamiltonian system in classical mechanics. Here the Legendre transformation from the Lagrangian to the Hamiltonian variables is singular (non-invertible). As a result the Hamiltonian variables are not all independent, but satisfy identities known as constraints. This in turn means that the Hamiltonian equations underdetermine the time-evolution of the Hamiltonian variables, leading to a gauge freedom in the description of the time-evolution, which means in other words a breakdown of determinism for the evolution of the state of the system as specified by the Hamiltonian variables.

More formally the arena for describing a constrained Hamiltonian system is what mathematicians call a *presymplectic manifold.* This is effectively a phase space equipped with a degenerate symplectic two-form ω. By degenerate one means that the equation $\omega(X) = 0$, where X is a tangent vector field, has non-trivial solutions, the integral curves of which we shall refer to as null curves on the phase space. The equations of motion are given in the usual Hamiltonian form as $\omega(X) = dH$, where H is the Hamiltonian function. The integral curves derived from this equation represent the dynamical trajectories in the phase space. But in the case we are considering there are many trajectories issuing from some initial point p_0, at time t_0. At a later time t the possible solutions of the Hamiltonian

[2]The treatment of this topic broadly follows the excellent account in Belot (1998).

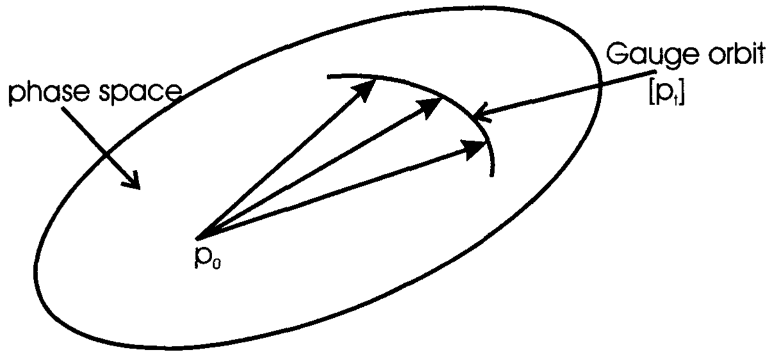


Fig. 14.5 *The indeterministic time evolution of a constrained Hamiltonian system.*

equations all lie on a *gauge orbit* in the phase space which is what we may call a null subspace of the phase space, in the sense that any two points on the orbit can be joined by a null curve as we have defined it. The situation is illustrated schematically in figure 14.5.

Instead of the initial phase point p_0 developing into a unique state p_t at a later time t as in the case of an unconstrained Hamiltonian system, we now have an indeterministic time-evolution, with a unique p_t replaced by a gauge orbit, which we denote by $[p_t]$ in figure 14.5. Effectively what is happening here is that the "physical" degrees of freedom at time t are being multiply represented by points on the gauge orbit $[p_t]$ at time t in terms of the "unphysical" degrees of freedom.

A familiar example of a constrained Hamiltonian system is the case of electromagnetism described by Maxwell's equations in vacuo. Here the Hamiltonian variables may be taken as the magnetic vector potential $\vec{A}$ and the electric field $\vec{E}$ subject to the constraint $div\ E = 0$. On a gauge orbit $\vec{E}$ is constant but $\vec{A}$ is specified only up to the gradient of a scalar function. The magnetic induction $\vec{B}$ defined by $\vec{B} = curl\ \vec{A}$ is then also gauge-invariant, i.e. constant on a gauge orbit. So $\vec{A}$ involves unphysical degrees of freedom, whose time-evolution is not uniquely determined. It is only for the physical degrees of freedom represented by $\vec{E}$ and $\vec{B}$ that determinism is restored. The gauge freedom in $\vec{A}$ belongs to surplus structure in the terminology of section 14.4.

14.6 Yang-Mills Gauge Theories

We turn now to a still more restricted sense of gauge symmetry associated with Yang-Mills gauge theories of particle interactions. To bring out the main idea we shall consider the simplest case of the non-relativistic (first-quantized) Schrödinger field. The field amplitude $\psi(x)$ (for simplicity we consider just one spatial dimension for the time being) is a complex number, but quantities like the charge density $\phi = e\psi^*\psi$ and the current density $j = \frac{1}{2}ie\left(\psi^*\frac{d\psi}{dx} - \psi\frac{d}{dx}\psi^*\right)$ are real quantities and can represent physical magnitudes. Consider now phase transformations of the form $\psi \rightarrow \psi e^{i\alpha}$. These are known as global gauge transformations since the phase factor α does not depend on x. If we now demand invariance of physical magnitudes under such gauge transformations, then ϕ and j satisfy this requirement. But suppose we impose *local* gauge invariance, i.e. allow the phase factor α to be a function $\alpha(x)$ of x. ϕ remains invariant but j does not. In order to obtain a gauge-invariant current we introduce the following device. Replace $\frac{d}{dx}$ by a new sort of derivative $\frac{d}{dx} - iA(x)$ where A transforms according to $A \rightarrow A + \frac{d}{dx}\alpha(x)$. Then the modified current $j(x) = \frac{1}{2}ie\left(\psi^*\left(\frac{d\psi}{dx} - iA\right) - \psi\left(\frac{d}{dx} + iA\right)\psi^*\right)$ *is* gauge-invariant. But this has been achieved by introducing a new field $A(x)$ as a necessary concomitant of the original field $\psi(x)$. Reverting to three spatial dimensions, the A field can be identified (modulo the electronic charge e) with the magnetic vector potential and the transformation law for $\vec{A}$ is exactly that described for the vector potential in the last section. The requirement of local gauge-invariance can be seen as requiring the introduction of a magnetic interaction for the ψ field.

Again we have an example here of physical structure being controlled by requirements imposed on surplus mathematical structure. The situation is illustrated schematically in figure 14.6. p_1, p_2, p_3 are three physical magnitudes, for example the charge or current at three different spatial locations. They are mapped onto m_1, m_2, m_3 in the mathematical structure M which is a substructure in the larger structure M'. The circles c_1, c_2, c_3 in the surplus structure represent possible phase angles associated with m_1, m_2, m_3 in a many-one fashion as represented by the arrows projecting c_1, c_2, c_3 onto m_1, m_2, m_3. Local gauge transformations represented by the arrows on the circles act independently at different spatial locations. They correspond to identity transformations on M and correlatively on P.

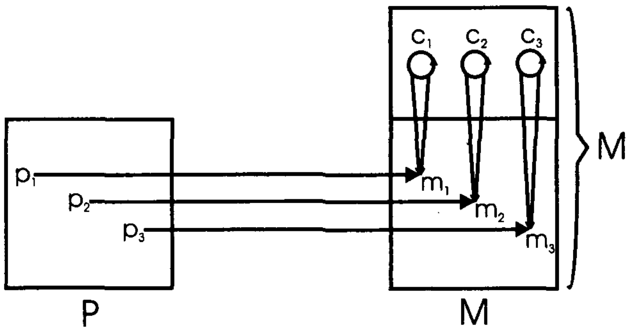


Fig. 14.6 *Gauge transformations and surplus structure.*

The $\vec{A}$ field establishes what mathematicians call a connection, correlating phases on the different circles c_1, c_2, c_3. The gauge transformations alter the connection as well as the individual phases in such a way as to maintain the gauge-invariance of the corrected "derivative" $\vec{\nabla} - i\vec{A}$.

Two ways of dealing with the surplus structure inherent in gauge theories suggest themselves. Firstly, we might just fix the gauge by some arbitrary convention,[3] but then we have lost the possibility of expressing gauge transformations which lead from one gauge to another. Alternatively, we might try to formulate the theory in terms of gauge-invariant quantities, which are the physically "real" quantities in the theory. Thus instead of the gauge potential, the $\vec{A}$ field in electromagnetism, we should employ the magnetic induction $\vec{B}$, specified by the equation $\vec{B} = curl\ \vec{A}$.

However, this manoeuvre has the serious disadvantage of rendering theory nonlocal! This is most clearly seen in the Aharanov-Bohm effect[4] in which a phase shift occurs between electron waves propagating above and below a long (in principle infinitely long) solenoid. The experiment is illustrated schematically in figure 14.7.

The magnetic induction is, of course, confined within the solenoid, so if it is regarded as responsible for the phase shift, it must be regarded as acting

[3]In some pathological cases this may not be consistently possible, a phenomenon known in the trade as the Gribov obstruction.

[4]The interpretation of the Aharonov-Bohm effect has occasioned considerable controversy in the philosophical literature. See, in particular, Healey (1997), Belot (1998) and Leeds (1999).

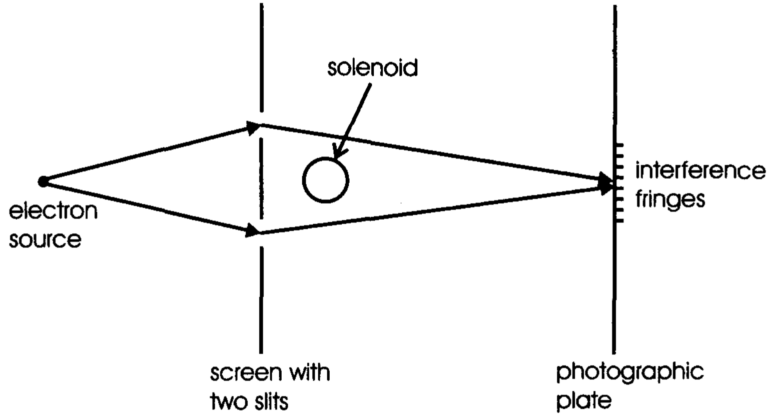


Fig. 14.7 *The Aharonov-Bohm experiment.*

nonlocally. On the other hand the vector potential extends everywhere outside the solenoid, so if invested with physical reality its effect on the electron phases can be understood as occurring locally. This is an argument for extending physical reality to elements which originated as elements of surplus structure.

However, just as in the case of free electromagnetism discussed in the previous section, the time-evolution of the vector potential is indeterministic since it is only specified up to the unfolding of an, in general, time-dependent gauge transformation. To restore determinism we must regard the gauge as being determined by additional "hidden variables" which pick out the One True Gauge, this seems a highly ad hoc way of proceeding as a remedy for restoring determinism. This is indeed a quite general feature of Yang-Mills gauge theories.[5]

14.7 The Case of General Relativity

The general arena for Yang-Mills gauge theories is provided by the notion of a fibre bundle. Speaking crudely a fibre bundle can be thought of as being constructed by attaching one sort of space, the fibre, to each point of

[5]For a detailed discussion see Lyre (1999).

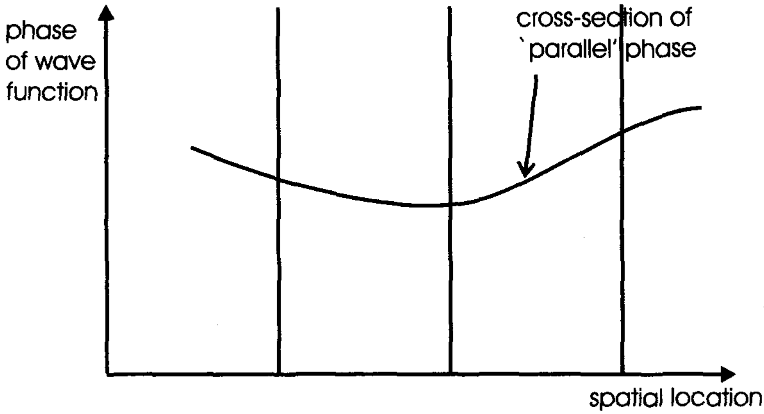


Fig. 14.8 *Fibre bundle structure of Yang-Mills gauge theory corresponding to figure 14.6.*

a second sort of space, the base space, so that *locally* the structure is just the familiar Cartesian product.

We can effectively redraw figure 14.6 in a way that brings out the bundle structure, as illustrated in figure 14.8. The local gauge group changes the phases according to the action of the $U(1)$ group. A cross-section of "parallel" or constant phase is specified by the connection field, i.e. the gauge potential.

In the case of general relativity (GR) we are dealing with the bundle of tangent spaces at each point of the spacetime manifold, or more appositely the frame bundle, specifying the basis (or frame) for the tangent space at every point. The gauge group is now the group of general 4-dimensional frame transformations, usually denoted by $GL(4,\mathbb{R})$. If consideration is restricted to Lorentzian frames the gauge group reduces to the familiar Lorentz group $SO(1,3)$ (or one might want to consider $SL(2,\mathbb{C})$, the covering group of $SO(1,3)$, if spinor fields are to be introduced). There are now two ways to go. Stick with the Lorentz group, and introduce a connection field to define parallel transport of frames from one point of spacetime to another. This was the original approach of Utiyama (1956). But it has been claimed repeatedly in the literature that if one wants to generalize classical relativity, so as to allow for torsion in the spacetime manifold, it is necessary to introduce an affine structure into the fibres (to be sharply distinguished from an affine connection on the bundle), so the local symmetry group becomes the *inhomogeneous* Lorentz group, i.e. the Poincaré group. Of course, this can be done from a purely mathematical point of view, but does not really make any *physical* sense at all. The translation subgroup effectively changes the origin, i.e. the point of attachment of the tangent space to the spacetime manifold, so inhomogeneous frame transformations correspond picturesquely to sliding the tangent space over the base space, but that is *not* what local gauge transformation are supposed to do—they move points around in the fibre at a *fixed* point on the base space. I refer the reader to Invanenko and Sardanashvily (1983) or Göckeler and Schücker (1987), who support, in my view correctly, the view that we do not need an affine bundle at all in order to extend GR to the Einstein-Cartan $\mathbb{U}_4$ theory incorporating spin and torsion.

So there is considerable confusion as between the Lorentz group and the Poincaré group as the appropriate Yang-Mills gauge group for GR and its generalizations, but it is also often claimed that general coordinate transformations (the subject of general covariance) provide the gauge group of GR! The following comments are intended to clarify what is going on here. Firstly, it should be noted that general coordinate transformations do not in general constitute a group from the global point of view, since in general they cannot be defined globally. But there is a globally defined symmetry group, which is an invariance group of GR, namely the diffeomorphism group, *diff*, which from the local point of view is the active version of local coordinate transformations. From the bundle point of view described above, elements of *diff* move points around in the base space, which is just the spacetime manifold. This is not directly connected with gauge freedom in the more specialized sense we have defined, that is to say either in the Yang-Mills sense or as arising in the theory of constrained Hamiltonian systems as described in section 14.5 above. To link up with the latter notion, we need to exhibit GR in a canonical formulation, sometimes referred to as the (3+1) approach to GR as compared with the 4-dimensional approach of the more familiar covariant formulation. In the (3+1) approach the configuration variables are the 3-geometries on a spatial slice at a given coordinate time. (The collection of all possible 3-geometries is what is often referred to as *superspace*.) The Hamiltonian (canonical) variables satisfy constraints, indeed the Hamiltonian itself vanishes identically. The gauge freedom arises essentially as a *manifestation* of the diffeomorphism invariance of the 4-dimensional covariant formulation, in the (3+1) setting. In this setting there are two sorts of gauge motion, one sort acting in the spatial slices and corresponding to diffeomorphisms of the 3-geometries, the other acting in time-like directions and corresponding to time-evolution of the 3-geometries.

The fact that time-evolution is a gauge motion, and hence does not correspond to any change at all in the "physical" degrees of freedom in the theory, produces the famous "problem of time" in canonical GR! Crudely this is often referred to under the slogan "time does not exist!" In a Pickwickian sense the indeterminism problem for constrained Hamiltonian systems is solved because time-evolution itself lies in a gauge orbit rather than cutting across gauge orbits, as in figure 14.5. The solution of the problem of time (which plagues attempts to quantize canonical GR), must involve in some way identifying some combination of the *physical* degrees of freedom with an *internal* time variable. But exactly how to do this remains a matter of controversy among the experts in canonical approaches to quantum gravity.[6]

14.8 The BRST Symmetry

In the path integral approach to general (non-Abelian) gauge theories, a naive approach would involve integrating over paths which are connected by gauge transformations. To make physical sense of the theory, the obvious move is to "fix the gauge", so that each path intersects each gauge orbit in just one point. However early attempts to derive Feynman rules for expanding the gauge-fixed path integral in a perturbation expansion led to an unexpected breakdown of unitarity.[7] This was dealt with in an ad hoc fashion by introducing fictitious fields, later termed ghost fields, which only circulated on internal lines of the Feynman diagrams in such a way as to cure the unitarity problem, but could never occur as real quanta propagating along the external lines of the diagrams. So getting rid of one sort of surplus structure, the unphysical gauge freedom, seemed to involve one in a new sort of surplus structure associated with the ghost fields.

[6] For a comprehensive account of canonical quantum gravity and the "problem of time" reference may be made to Isham (1993).

[7] Cp. Feynman (1963).

The whole situation was greatly clarified by the work of Fadeev and Popov (1967) who pointed out that when fixing the gauge in the path integral careful consideration must be given to transforming the measure over the paths appropriately. The transformation of the measure was expressed in a purely mathematical manoeuvre as an integral over scalar Grassmann (i.e. anticommuting) fields which were none other than the ghost (and antighost) fields!

The effective Lagrangian density could now be written as the sum of three terms, $\mathcal{L}_{eff} = \mathcal{L}_{gi} + \mathcal{L}_{gf} + \mathcal{L}_{ghost}$, where $\mathcal{L}_{gi}$ is a gauge-invariant part, $\mathcal{L}_{gf}$ is a non-gauge-invariant part arising from the gauge fixing, and $\mathcal{L}_{ghost}$ is the contribution from the ghost fields.

$\mathcal{L}_{eff}$ no longer, of course, has the property of gauge invariance, but it was discovered by Becchi, Rouet and Stora (1975) and independently by Tyutin (1975) that $\mathcal{L}_{eff}$ does exhibit a kind of generalized gauge symmetry, now known as BRST symmetry, in which the non-invariance of $\mathcal{L}_{gf}$ is compensated by a suitable transformation of the ghost fields contributing to $\mathcal{L}_{ghost}$.

To see how this comes about we consider the simplest (Abelian) case of scalar electrodynamics. The matter field ψ satisfies the familiar Klein-Gordon equation. Under the local gauge transformation $\psi \rightarrow \psi e^{i\alpha(x)}$, where x now stands for the 4-dimensional spacetime location x^μ, the gauge-invariance of the Lagrangian for the free field is restored by using the corrected derivative $\partial \rightarrow \partial_\mu - iA_\mu$, where the gauge potential A_μ can be identified, modulo the electronic charge, with the electromagnetic 4-potential. A_μ transforms as $A_\mu \rightarrow A_\mu + \partial_\mu \alpha(x)$. The field strength $F_{\mu\nu} = A_{\nu,\mu} - A_{\mu,\nu}$ is gauge-invariant and measures the curvature of the connection field A_μ in the geometrical fibre bundle language. All that we have done here is just a relativistic generalization of the discussion already given in section 14.6.

To formulate the BRST transformation we consider a 5-component object

$$\Phi = \begin{pmatrix} \psi \\ A_\mu \\ \eta \\ \omega \\ b \end{pmatrix}$$

where ψ is the matter field, A_μ the gauge potential which we have already introduced above, η is the ghost field, ω the antighost field, and b is what is usually termed a Nakanishi-Lautrup field.

η and ω are anticommuting (Grassmann) scalar fields. The fact that they violate the spin-statistic theorem, which would associate scalar fields with commuting variables, emphasizes the unphysical character of the ghosts and antighosts.

We have then

$$\omega^2 = \eta^2 = 0.$$

The BRST symmetry is defined by

$$\Phi \longrightarrow \Phi + \epsilon s\Phi$$

where ϵ is an infinitesimal Grassmann parameter and

$$s\Phi = \begin{pmatrix} i\eta\psi \\ \partial_\mu\eta \\ 0 \\ b \\ 0 \end{pmatrix}$$

The first two components of $s\Phi$ comprise just the infinitesimal version of a gauge transformation with the arbitrary spacetime function $\alpha(x)$ replaced by the ghost field η. But ϵ is a constant so the BRST transformation is a curious hybrid. It is in essence a nonlinear rigid fermionic transformation, which contains within itself, so to speak, a local gauge transformation specified by a dynamical field, namely the ghost field.

What is the role of the Nakanishi-Lautrup field? By incorporating this field the transformation is rendered nilpotent,[8] i.e. it is easily checked that $s^2\Phi = 0$. But this means that s behaves like an exterior derivative on the extended space of fields. This in turn leads to a beautiful generalized de Rham cohomology theory in terms of which delicate properties of gauge fields, such as the presence of anomalies, the violation of a classically imposed symmetry in the quantized version of the theory, can be given an elegant geometrical interpretation.[9]

But now we can go further. Instead of arriving at the BRST symmetry via the Fadeev-Popov formalism, we can forget all about gauge symmetry in the original Yang-Mills sense, and impose BRST symmetry directly as

[8] The original BRST transformation failed to be nilpotent on the antighost sector.

[9] See Fine and Fine (1997) for an excellent account of these developments.

the fundamental symmetry principle. It turns out that this is all that is required to prove the renormalizability of anomaly-free gauge theories such as those considered in the standard model of the strong and electroweak interactions of the elementary particles.

But we may note in passing that for still more recondite gauge theories further generalizations have had to be introduced.[10]

(1) In a sense the ghosts compensate for the unphysical degrees of freedom in the original gauge theories. But in some cases the ghosts can "over-compensate" and this has to be corrected by introducing ghosts of ghosts, and indeed ghosts of ghosts of ghosts etc.!
(2) For the more general actions contemplated in string and membrane theories the so-called Batalin-Vilkovisky antifield formalism has been developed. This introduces partners (antifields) for all the fields, but the antifield of a ghost is not an antighost and the anti (antighost) is not a ghost!

14.9 Conclusion

As we have seen, there are three main approaches to interpreting the gauge potentials.

The first is to try and invest them with physical reality, i.e. to move them across the boundary from surplus structure to M in the language of figure 14.4. The advantage is that we may then be able to tell a local story as to how the gauge potentials bring about the relative phase shifts between the electron wave functions in the Aharanov-Bohm effect, but the disadvantage is that the theory becomes indeterministic unless we introduce ad hoc hidden variables that pick out the One True Gauge.

The second approach is to try and reformulate the whole theory in terms of gauge-invariant quantities. But then the theory becomes nonlocal. In the case of the Aharonov-Bohm effect this can be seen in two ways. If the phase shift is attributed to the gauge-invariant magnetic induction this is confined *within* the solenoid whereas the experiment is designed so that the electron waves propagate *outside* the solenoid. Alternatively we might try to interpret the effect not in terms of the $\vec{A}$ field itself which of course is

[10] Weinberg (1996), Chapter 15, may be consulted for further information on these matters.

not gauge-invariant but in terms of the gauge-invariant holonomy integral $\oint \vec{A}\cdot d\vec{l}$ taken round a closed curve C encircling the solenoid. (This by Stokes theorem is of course just equal to the flux of magnetic induction through the solenoid.) But if the fundamental physical quantities are holonomies, then the theory is again clearly "nonlocal", since these holonomies are functions defined on a space of loops, rather than a space of points.

Furthermore, with this second approach, the principle of gauge invariance cannot even be formulated since gauge transformations are defined by their action on non-gauge-invariant quantities such as gauge potentials, and in the approach we are now considering the idea is to eschew the introduction of non-gauge-invariant quantities altogether!

So this leaves us with the third approach. Allow non-gauge-invariant quantities to enter the theory via surplus structure. And then develop the theory by introducing still more surplus structure, such as ghost fields, antifields and so on. This is the route that has actually been followed in the practical development of the concept of gauge symmetry as we have described in the previous section.

But this leaves us with a mysterious, even mystical, Platonist-Pythagorean role for purely mathematical considerations in theoretical physics. This is a situation which is quite congenial to most practising physicists. But it is something which philosophers have probably not paid sufficient attention to in discussing the foundations of physics. The gauge principle is generally regarded as the most fundamental cornerstone of modern theoretical physics. In my view its elucidation is the most pressing problem in current philosophy of physics. The aim of the present paper has been, not so much to provide solutions, but rather to lay out the options that need to be discussed, in as clear a fashion as possible.

References

Becchi, C., Rouet, A. and Stora, R. (1975): Renormalization of the Abelian Higgs-Kibble model. *Communications in Mathematical Physics*, 42:127–162.

Belot, G. (1998): Understanding electromagnetism. *The British Journal for the Philosophy of Science*, 49:531–555.

Fadeev, L.D. and Popov, V.N. (1967): Feynman diagrams for the Yang-Mills field. *Physics Letters B*, 25:29–30.

Feynman, R.P. (1963): Quantum theory of gravity. *Acta Physica Polonica*, 24:697–722.

Fine, D. and Fine, A. (1997): Gauge theory, anomalies and global geometry: The interplay of physics and mathematics. *Studies in History and Philosophy of Modern Physics*, 28:307–323.

Göckeler, M. and Schücker, T. (1987): *Differential Geometry, Gauge Theories, and Gravity.* Cambridge: Cambridge University Press.

Healey, R. (1997): Nonlocality and the Aharonov-Bohm effect, *Philosophy of Science*, 64:18–40.

Isham, C.J. (1993): Canonical quantum gravity and the problem of time. In L.A. Ibort and M.A. Rodriguez (eds.), *Integrable Systems, Quantum Groups, and Quantum Field Theories.* Dordrecht: Kluwer, pp.157–287.

Ivanenko, D. and Sardanashvily, G. (1983): The gauge treatment of gravity. *Physics Reports*, 94:1–45.

Leeds, S. (1999): Gauges: Aharonov, Bohm, Yang, Healey. *Philosophy of Science*, 66:606–627.

Lyre, H. (1999): Gauges, holes and their 'connections'. *Fifth International Conference on the History and Foundations of General Relativity.* E-print gr-qc/9904036.

Pauli, W. (1941): Relativistic field theories of elementary particles. *Reviews of Modern Physics*, 13:203-232.

Redhead, M.L.G. (2001): The intelligibility of the universe. In A. O'Hear (ed.), *Philosophy at the New Millennium.* Cambridge: Cambridge University Press, pp. 73-90.

Shaw, R. (1954): *The problem of particle types and other contributions to the theory of elementary particles.* Cambridge University PhD thesis.

Tyutin, I.V. (1975): Gauge invariance in field theory and statistical mechanics. *Lebedev Institute Reprint N 39.*

Utiyama, R. (1956): Invariant Theoretical Interpretation of Interaction. *Physical Review*, 101:1597-1607.

Weinberg, S. (1996): *The Quantum Theory of Fields.* Vol. II, *Modern Applications.* Cambridge: Cambridge University Press.

Weyl, H. (1918): Gravitation und Elektrizität. *Sitzungsberichte der Preussischen Akademie der Wissenschaften.* 465-480.

Weyl, H. (1929): Elektron und Gravitation. *Zeitschrift für Physik*, 56:330-352.

Yang, C.N. and Mills, R.L. (1954): Conservation of isotopic spin and isotopic gauge invariance. *Physical Review*, 96:191-195.

This page is intentionally left blank

Chapter 15

Comment on Redhead: The Interpretation of Gauge Symmetry

Michael Drieschner
Ruhr-University Bochum

Tim Oliver Eynck
NIKHEF, Amsterdam

Holger Lyre
University of Bonn

Abstract. *In his contribution to these proceedings Michael Redhead investigates the concept of gauge symmetry in contemporary physics, an important endeavor in view of the current mismatch between the scant attention this topic has so far received from philosophers of science and its key role in the description of at least three of the fundamental forces known today.*

After a short historical account Redhead relates the term gauge to a general ambiguity inherent to any mathematical description of physical phenomena. Constrained Hamiltonian systems in general, and Yang-Mills theories in particular, constitute examples of physical models necessitating the introduction of mathematical surplus structure. *The latter arises when dealing with a mathematical structure larger than strictly necessary to represent a given physical structure (cf. also Redhead 1975, 1998, 2001). Redhead includes in this account even recent developments such as BRST symmetry. His presentation leads to an interpretational trilemma elucidating the ontological status of gauge fields.*

Three points from Redhead's analysis have been chosen to be commented on (semi-) independently in the following. Section 15.1 discusses the relation of the general concept of ambiguity in mathematical description to the particular case of gauge symmetry. Section 15.2 introduces a notion of "prepotentials" *as a tool for the investigation of questions concerning the ontology and locality of gauge potentials.*

Section 15.3 mainly deals with possible solutions to Redhead's trilemma. We end with some concluding remarks.

15.1 The Ubiquitous Ambiguity

As pointed out already, Redhead's discussion of gauge symmetry is embedded in a more general one, pertaining to the ambiguities in the mathematical representation of any physical phenomenon. The plural here has been chosen deliberately. We find various ambiguities (e.g. in a scale of mineral scratchability), but the analysis of those connected to symmetries of the physical system under consideration seems most fruitful to us. We believe it can reveal at least part of the nature of this peculiar surplus structure that arises whenever the physics is not surjectively mapped onto the mathematics. Examples for this are provided by constrained Hamiltonian systems (already at the level of classical mechanics). In field theories local gauge-invariance seems to be the only ingredient, provided one holds no objections against the ambiguous gauge potential needed to derive an interacting theory. But if gauge symmetry provides but one example of an ambiguity related to a physical symmetry, why does it deserve special attention?

Practical importance in physics cannot be the only reason, for then one could carry out case studies on less complicated systems and simply translate all results obtained to gauge language afterwards. An important motivation could be the non-local features of gauge theories. To what extent locality is jeopardized or preserved by shifting the border between physical and surplus structure will be discussed in section 15.3. At this stage we prefer to draw attention to the relationship between symmetry and geometry. In pure mathematics the concept of investigating geometrical structures by means of symmetry groups has a long and successful history (e.g. the Erlangen programme). It seems hardly surprising that this has repercussions in the science that utilizes in its description of dynamics in space and time the tools mathematicians have developed in *their* study of spatial structures. In other words, both general relativity and the fibre bundle formulation of gauge theories are prime examples for a *geometrization of physics*.

But even our maturing formulations of classical mechanics and their ontological analysis can serve to illustrate this tendency: manifolds come in their very definition equipped with a diffeomorphism invariance—they intrinsically favour neither a substantival nor a relational interpretation of space and time. Viewed in this light, gauge theories appear as but a sophisticated version of this geometrization, which now comprises besides external also internal spatial structures, ascending in complexity from differentiable or metric structures to such refined notions as cohomology. These latter concepts may turn out to be the only ones appropriate to state precisely which spatial (geometrical) properties are ascribed ontological relevance in physical theories. In a first-quantized fibre bundle formulation these could be certain equivalence classes of connections (cf. section 15.2) or BRST cohomology classes in second field quantization, but, to be sure, in none of these cases any particular representatives. In this way, the geometrization of modern physics provides a viable albeit formal and barely intuitive answer to the question of surplus structure. To put it more explicitly, the formulation of dynamics on a space of equivalence classes is in no way particular to gauge theories. The ambiguous choice of representatives stemming from this should not be confused with a fundamental indeterminism.

However, gauge theories do show at least one novel feature relevant in this context, namely the Aharanov-Bohm effect (Aharonov and Bohm 1959)—for short: AB effect. It relates the gauge freedom to issues of locality (cf. sections 15.3 and 15.4). Contrary to the central role the AB effect does and will play in this comment, Redhead seems to attach less importance to it in his analysis of the alleged gauge-related indeterminism, an aspect of the discussion which, from our point of view, deserves further attention.

15.2 "Prepotentials"

What makes the well-known AB effect so astonishing is that there seems to exist "nothing" in the region of space where the electron's possible paths extend, yet the electron is discernibly influenced by "something". Thus apparently physics does not only "locally" depend on the interaction fields: if loop integrals enter the game—around a region that contains non-zero field strengths—then the fields inside the solenoid have an effect also outside it.

Before proceeding with our analysis of the AB effect we should to some extent explicate the terms enclosed by inverted commas in the previous paragraph. We will in the following be dealing with the *ontology*—questions concerning the existence—of theoretical entities, in particular gauge potentials. Far from attempting an exhaustive definition we propose as one necessary criterion for the reality of such a theoretical entity that its alteration have a discernible (physical) effect. Likewise, we shall refer to questions concerning the relation between the spatial situation of physical entities and their potentiality to interact as pertaining to *locality.*

In classical electrodynamics only the fields represent something physically real, whereas by carrying out (smooth) gauge transformations we can arbitrarily change the potentials, which hence cannot be real. Were it purely classical, the AB effect would therefore teach us that charged particles interact non-locally with the electromagnetic field. If in a quantum theory the wave function of the electron coupled to the field strength, the same conclusion would hold there. To avoid this consequence it has been suggested that maybe the ψ-function of the electron interacts locally "far outside" the solenoid where the magnetic field lines close. But this solution has to be repudiated since one can enclose the magnetic field to a very high precision within toroidal solenoids and does not observe any dependence of the effect on the quality of the shielding (cf. Peshkin and Tonomura 1989, part 2 for a detailed discussion).

Thus the description in terms of fields alone as contemplated above turns out to be incomplete: Even regions of space with vanishing electromagnetic fields nevertheless admit *different* configurations of electromagnetic potentials that cannot be transformed into each other by gauge transformations. For example, in the AB setting the potential cannot be gauge transformed to zero everywhere outside the solenoid. This fact suggests a formulation of the surplus structure that might lend itself to analyzing ontology and locality of the entities under consideration.

Again, let us interrupt our discussion of the AB effect, the particular case at hand, to briefly lay out in general terms what we are aiming at. Whenever confronted with an ambiguous map from the physical to some mathematical objects, it seems natural to try to restrict the set of images to equivalence classes. This procedure is well known, e.g. from propositional logic: One begins with statements that may be logically equivalent (imply each other) even though they differ, for instance in language, as "The hat is red" and "Der Hut ist rot". To remove this ambiguity one passes over to equivalence classes of statements, usually called *propositions.* The two statements quoted above are then representatives of the same proposition, and one can rightly state the basic law of logic *(anti-symmetry)*: "Propositions that imply each other are equal". Obviously, this procedure is very general indeed, although for a particular theory under study one has to find a suitable equivalence relation, possibly a highly non-trivial task.

Recall that the basis of our considerations is that electromagnetic potentials are equivalent if they differ only by a divergence: $A'_\mu \sim A_\mu$ iff there is a (sufficiently differentiable) $\alpha(x)$ such that $A'_\mu = A_\mu - \partial_\mu \alpha(x)$. Or, to put it differently, iff they can be transformed into each other by a "gauge" transformation. Thus we can introduce *equivalence classes of electromagnetic potentials*, their equivalence relation being that they can be related by gauge transformations well-defined *everywhere on the base space*. One can easily see that this indeed qualifies as an equivalence relation since gauge transformations form a group, the gauge group. Certainly, an equivalence class of potentials constitutes an entity much more abstract than a field, as the result of the strategy outlined in the previous paragraph one could view it as a sort of *"prepotential"*. It will in general not be possible, for instance, to ascribe to such a prepotential a *value*, but quite to the contrary, every equivalence class contains potentials with *any* value at some given spacetime point. The construction presented here is, of course, mathematically equivalent to the so-called *loop approach* (based on the pioneering work of Yang 1974 and Wu and Yang 1975, for a very accessible recent account cf. Gambini and Pullin 1996) and in the AB setting reflects the non-trivial topology of the base space. In this respect the very concept of a prepotential is non-local—or rather, to be precise, non-separable.[1]

This formulation shares with the more common one in terms of gauge potentials the property that we can choose any electromagnetic potential from the prepotential equivalence class in order to do calculations. The condition that all our calculations be gauge invariant implies that they have to be valid for all representatives of the prepotential, it thus exemplifies the general structures introduced in the previous paragraph as well as those discussed in section 15.1.

What does that mean for our questions of ontology and locality? Let us illustrate our approach with the AB effect: From the experimental setup—a magnetic field, spatially strictly enclosed within a toroid—it looks at first glance as if there were nothing outside the solenoid. In fact, the electro-

[1]In section 15.3 we will briefly comment on the conceptual distinction between non-locality and non-separability. Since Redhead himself does not make such a distinction we will mostly use the term "non-local" in both senses—as he does.

magnetic fields are zero everywhere in the outside space. The prepotential formulation gives us now a possibility to see more clearly what happens. There is something outside the toroid that we may call real because it has a physical effect: the prepotential! It is physically different from the one containing the potential "$A_\mu(x) \equiv 0$ everywhere in the outside region". And this is a "real" difference, since there is no surplus structure in those prepotentials any more that could be considered non-real. So the prepotentials meet our criteria laid out in this section and hence qualify as real entities.

As far as its basic entities are concerned, the description presented here is, as mentioned earlier, completely equivalent to the loop space approach. Prepotentials or loops are, indeed, to be considered the basic entities of electromagnetic reality. One can even formulate the full dynamics of any (pure) gauge theory on loop space (cf. again Gambini and Pullin 1996), an aspect that did not concern us here. Perhaps our formulation in terms of prepotentials has the didactic merit of giving us a somewhat more intuitive picture and of placing the loop approach into a more general context.

15.3 Redhead's Trilemma

In its concluding section, Redhead's analysis culminates in a trilemma of the three main approaches to interpreting gauge potentials. Firstly, we may consider gauge potentials as real. *Prima facie*, this provides us with a local account of gauge theories—particularly in the case of the AB effect. The disadvantage is that we get an indeterministic influence from non-observable physical beables to observable ones (e.g. interference fringes). Hence, this option leaves us with a version of the notorious *hole argument* (Earman and Norton 1987).

Now, secondly, we may consider holonomies as real. This way we will get rid of the troublesome surplus structure by focusing on gauge-invariant quantities only, such as loop integrals $\oint A_\mu dx^\mu$ (i.e. holonomies). This, however, will render the theory non-local, since, as Redhead points out, "... *holonomies are functions defined on a space of loops, rather than a space of points*" (p. 299). In the example of the AB effect there is no pointlike local interaction between the magnetic field strength inside and the electron wave function outside the solenoid. Moreover, Redhead stresses the point that, apparently, we lose the ability of formulating the gauge principle, since local gauge transformations act on non-gauge-invariant quantities—such as gauge potentials and wave functions.

Finally, we may add even more surplus structure in terms of ghost fields, etc. There is some indication for this option with regard to its pragmatic success—at least for the practising physicist. However, the ontological price we pay is to allow for a "Platonist-Pythagorean" role for purely mathematical elements to influence physical elements of reality. This is the third horn of the trilemma. In brief, Redhead's trilemma looks like this:

(1) Consider gauge potentials as real.
Pro: local account
Con: indeterministic influence from non-observable physical beables to observable ones
(2) Consider holonomies as real.
Pro: no surplus structure, solely gauge-invariant quantities
Con: non-local account, renunciation of the gauge principle
(3) Add even more surplus structure.
Pro: local account, pragmatic success
Con: "Platonist" influence of mathematical elements on elements of reality

Our criticism of this trilemma starts with the remark that neither option 1 nor option 3 are purely local. Certainly, the sore point here is the term "local". Admittedly, the reality of gauge potentials provides us with a local account with pointlike field interactions (i.e. the gauge potential and electron wavefunction outside the solenoid). However, once we understand the term "local" in the sense of *local separability*, we are no longer able to consider the AB effect as "local". It is not a definite value of the gauge potential which is responsible for the effect, but rather the prepotential which leads to the same holonomy. This indicates the deep topological nature of the AB effect—stemming from the topology of the gauge group $U(1)$.[2] It is, therefore, impossible to tell a local story of the AB effect in terms of separable quantities (or properties), since, as Richard Healey puts it, "*... these properties do not supervene on any assignment of qualitative intrinsic physical properties at spacetime points in the region concerned*" (Healey 1997). Another example of non-separability is provided by the EPR correlations in quantum mechanics. Note, however, their quite different origin

[2]To be precise: the non-trivial mapping $\mathbb{S}^1 \to \mathbb{S}^1$ between configuration space and gauge group.

in quantum non-locality rather than topology—a more rigorous analysis of this will be given in a forthcoming paper (Eynck et al. 2001).

To put it differently: since holonomies do not uniquely correspond to regions of space, they render gauge theories non-separable (i.e. non-local in a particular sense). It is exactly this feature which counts as a disadvantage—in Redhead's view—in option 2. However, since none of the options may explain topological effects in gauge theories without referring to something like prepotentials or holonomies, the alleged advantage of "locality" of options 1 and 3 disappears.

15.4 Conclusion

Redhead tries to tentatively solve his own trilemma. He argues that option 3 has to be favoured, because—compared to option 2—it does retain the gauge principle. This latter claim seems to require further clarification since it appears somewhat unclear which feature of the loop approach could possibly prevent one from picking a particular gauge potential from the many equivalent ones and performing with this representative all the usual manipulations advertised in the textbooks. Besides, as has been shown by several authors[3], the gauge principle is in any case not sufficient to derive the interaction-coupling of gauge theories.

Therefore, in our view, option 2 seems to be *the* preferable candidate. In section 15.2 we gave a description in terms of prepotentials, i.e. non-separable equivalence classes of gauge potentials in the whole of space—an account that has some pedagogical merit but is otherwise equivalent to the loop integral formulation and, hence, Redhead's option 2. One should not be bothered by the non-separable features of this interpretation, since they are shared by option 3.

However, the conclusion, which goes right to the heart of the matter, is that the seemingly problematic surplus structure features sometimes implicitly but nevertheless omnipresently in modern physics, as we already argued in more general terms in the beginning section. Physics is thus essentially always limited to formulations up to isomorphisms and equivalence classes.

[3]Cf. Brown 1999, Healey 2001, Teller 2000 and Lyre 2000, 2001 (the last two references also propose a generalized equivalence principle providing the true empirical input for gauge theories).

This observation reflects the importance of group theory in modern fundamental accounts. Any mathematical description of nature as we know it seems to require the existence and interplay of both dynamical and invariant quantities, with (Lie) groups classifying and mediating between them. One could then wonder whether this sheds any light on such fundamental questions as why physics is possible at all, but pursuing this would certainly lead us too far afield.

Acknowledgements

Many thanks to the organizers, in particular to Manfred Stöckler and Andreas Bartels, who took on the bulk of the organizational preparations in the run-up to the conference.

References

Aharonov, Y. and Bohm, D. (1959): Significance of electromagnetic potentials in the quantum theory. *Physical Review*, 115(3):485–491.

Brown, H. R. (1999): Aspects of objectivity in quantum mechanics. In Butterfield, J. and Pagonis, C. (eds.), *From Physics to Philosophy.* Cambridge: Cambridge University Press.

Earman, J. and Norton, J. (1987): What price spacetime substantivalism? The hole story. *The British Journal for the Philosophy of Science*, 83:515–525.

Eynck, T. O., Lyre, H., and Rummell, N. v. (2001): A versus B! Topological nonseparability and the Aharonov-Bohm effect. *Quantum Structures V.* E-print PITT-PHIL-SCI00000404.

Gambini, R. and Pullin, J. (1996): *Loops, Knots, Gauge Theories and Quantum Gravity.* Cambridge: Cambridge University Press.

Healey, R. (1997): Nonlocality and the Aharonov-Bohm effect. *Philosophy of Science*, 64:18–41.

Healey, R. (2001): On the reality of gauge potentials. *Philosophy of Science*, 68(4):432–455.

Lyre, H. (2000): A generalized equivalence principle. *International Journal of Modern Physics D*, 9(6):633–647.

Lyre, H. (2001): The principles of gauging. *Philosophy of Science*, 68:S371–S381.

Peshkin, M. A. and Tonomura, A. (1989): *The Aharonov-Bohm Effect.* Lecture Notes in Physics 340. Berlin: Springer.

Redhead, M. L. G. (1975): Symmetry in Intertheory Relations, *Synthese*, 32:77–112.

Redhead, M. L. G. (1998): Review: S. Y. Auyang, How is quantum field theory possible? (New York, Oxford University Press, 1995). *The British Journal for the Philosophy of Science*, 49:499–507.

Redhead, M. L. G. (2001): The intelligibility of the universe. In A. O'Hear (ed), *Philosophy at the New Millennium.* Cambridge: Cambridge University Press, pp. 73-90.

Redhead, M. L. G. (2002): The interpretation of gauge symmetry. This volume.

Teller, P. (2000): The gauge argument. *Philosophy of Science*, 67 (Proceedings):S466–S481.

Wu, T. T. and Yang, C. N. (1975): Concept of nonintegrable phase factors and global formulation of gauge fields. *Physical Review D*, 12(12):3845–3857.

Yang, C. N. (1974): Integral formalism for gauge fields. *Physical Review Letters*, 33(7):445–447.

Chapter 16

Is the Zero-Point Energy Real?

Simon Saunders
Oxford University

Abstract. *I consider arguments to show that the vacuum energy density should receive a large contribution from the zero-point energy. This is the cosmological constant problem, as it was originally framed. The problem is compared to others that attended the notion of vacuum, in classical electromagnetism and in the Dirac hole theory. In both cases their resolution depended in part on a change in interpretation of extant theory , in part on new theory. I suggest a shift in the interpretation of quantum theory that does appear to have a bearing on the cosmological constant problem.*

16.1 Introduction

The nature of the vacuum state has proved to be of enduring interest in field theory. Time and again it has been refashioned. Evidently we need the right physical concepts, even to understand the state in which nothing exists.

When it comes to the vacuum of QFT it is far from clear that we have the correct physical concepts. Of course, on moving to the Planck scale there are plenty of reasons to question the principles of QFT, but the problem I have in mind, whilst it does concern gravity, arises at all scales. It can be posed as a problem of elementary quantum theory. The difficulty is this: it appears that there must remain a very large energy in the vacuum state, and that this should contribute massively to the source terms of general relativity. This problem has traditionally been called *the cosmological constant problem*, since the source term is proportional to the metric tensor, and hence should yield a very large cosmological constant, wildly in excess of the observed value. The problem has also been generalized; there are other mechanisms whereby the vacuum appears to pick up a large energy, again in conflict with observation. As such it has received a great deal of attention in recent years, but there is little consensus on how it may be solved.

Here I am concerned with the original statement of the problem, in terms of the zero-point energy of QFT. Is the zero-point energy real? The usual argument given for its reality is the Casimir effect; I shall consider this in detail in due course. But prior to that, I will attempt to gain some historical perspective on the problem. The vacuum of field theory has seen some radical changes, first, with the elimination of ether in classical electromagnetism, and second, with the elimination of the Dirac negative energy sea in quantum electrodynamics. The latter is particularly instructive; the negative energy of the Dirac vacuum can be viewed as the fermion zero-point energy by another name. Both examples are cases where the vacuum turned out not to have the problematic feature it was thought to have.

The question arises as to whether a similar fate awaits the zero-point energy—whether in fact the cosmological constant problem (as traditionally formulated) is a spurious one. But if so, and if the previous historical examples are anything to go by, not just a change in philosophy is needed; there will be a change in physics as well. I will at the close suggest a change in philosophy, but not yet a change in the physics. My suggestion is doubly limited, in that it has no bearing on the other ways in which the vacuum state can pick up energy—in particular the energy shifts, possibly large, that are expected to arise on spontaneous symmetry breaking.

16.2 The Cosmological Constant Problem

A common statement of the problem is as follows:

> The cosmological constant problem is one of the most serious puzzles confronting particle physics and cosmology. No symmetries or principles of General Relativity prohibit a cosmological constant term from appearing in the Einstein equations. Moreover, any vacuum energy such as that predicted by quantum field theory must—according to the equivalence principle—gravitate, and will act as a cosmological constant. However, the upper bound on the present

day cosmological constant is extremely small in particle physics units: $\frac{\lambda}{m^4_{Planck}} < 10^{-122}$ (Brandenburger 1999).

As stated it is a fine-tuning problem, with a bound that would be even smaller if we did not introduce a cut-off at the Planck scale (on the optimistic assumption that whatever physics comes into play in the Planck regime, it will not add to the vacuum energy). For a cutoff Λ, the zero-point energy of a field of mass $m \ll \Lambda$ is (with $h = c = 1$):

$$\int_0^\Lambda \frac{1}{2}\sqrt{k^2 + m^2}\, \frac{4\pi k^2 dk}{(2\pi)^3} \sim \frac{\Lambda^4}{16\pi^2}.$$

This quantity is just the sum of the zero-point energy over the normal modes of the field up to the cut-off Λ. If this is set at the Planck mass, $\Lambda \sim m_{Planck} \sim 10^{19}$ GeV, then given the current upper bound on the cosmological constant $\lambda < 10^{-29} g/cm^3 \sim (10^{-12}\text{ GeV})^4$, the observed value is about 122 orders of magnitude smaller than we expect. If the contribution from the zero-point energy is to be cancelled by the true cosmological constant, the latter will have to be equal to it and of opposite sign to one part in 10^{122}—making it the most accurately known number in physics.

Brandenberger goes on to suggest a mechanism whereby scalar gravitational fluctuations, with wavelength greater than the Hubble radius, are formed as a back-reaction to the presence of cosmological perturbations, which act as a negative cosmological constant in the de Sitter background. He suggests this mechanism may in fact be self-regulating, leading, more or less independent of the original value of the cosmological constant, to an effective value to it of order unity (on the Planck scale), which cancels the stress-energy tensor due to the zero-point energy.

This proposed solution is typical of the genre. Coleman's well-known proposal is the same (Coleman 1988): the cosmological constant becomes a dynamical variable in a certain Euclidean path-integral formulation of quantum gravity, whereby the amplitude is shown to be greatly peaked at a net value close to zero. Cancellation of the cosmological constant, with the source term due to the zero-point energy, is the name of the game. On Coleman's proposal, wormholes, connecting geometries, make the Euclidean action very large for geometries with net non-zero cosmological constants. They therefore make vanishingly small contribution to the path integral.

More recent attempts have considered quintessence, anthropic, k-essence, braneworld, and holographic approaches. The literature is large and still rapidly growing. Very little of it considers the original motivation for the problem critically—whether because the arguments are so clear-cut as to be unanswerable, or because here is a problem worth taking seriously just because there are lots of interesting conjectures that can be made about it. Very likely it is a mixture of the two:

> Physics thrives on crisis....Unfortunately, we have run short of crises lately. The 'standard model' of electroweak and strong interactions currently faces neither internal inconsistencies nor conflicts with experiment. It has plenty of loose ends; we know no reason why the quarks and leptons should have the masses they have, but then we know no reason why they should not.
>
> Perhaps it is for want of other crises to worry about that interest is increasingly centered on one veritable crisis: theoretical expectations for the cosmological constant exceed observational limits by some 120 orders of magnitude. (Weinberg 1989, p.1).

No doubt many would be disappointed if there turns out to be no good reason, after all, to take the zero-point energy seriously.

Here is the argument to show why the vacuum, if it contributes any energy at all, will yield an effective cosmological constant. By the Equivalence Principle, the local physics is Lorentz invariant, so in the absence of any local matter or radiation, locally we should see the same physics as in the vacuum. But the vacuum expectation value of the Minknowski space stress-energy tensor must be a multiple of the Minkowski metric. Therefore, we expect to find a term $\lambda g^{\mu\nu}$ on the RHS of the Einstein field equations. Such a term characterizes a perfect fluid with equation of state:

$$P_{vac} = -\rho_{vac}. \tag{16.1}$$

Under an adiabatic expansion from V to $V + dV$, an amount of work PdV is done, which provides exactly the mass-energy to fill the new volume $V + dV$ with the same energy-density ρ_{vac}. Expanding or compressing nothing changes nothing, as one would expect. Since locally we expect the equation

$$R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R = \frac{8\pi G}{c^4}\langle T^{\mu\nu}\rangle_\omega \tag{16.2}$$

to hold for the observed metric and curvature (the c-number quantities on the LHS), then in a local vacuum state ω we expect the RHS to contribute a term $\lambda g^{\mu\nu}$—and, indeed, if ω is anything like the Fock space vacuum, with a formally divergent value for λ, corresponding to a zero-point energy $\frac{1}{2}h\nu$ for each normal mode of a quantum field.

Of course the notion of "local vacuum state" is not easily defined in QFT. We cannot use number density operators to define it, for these are non-local quantities. But one would expect that the Fock space vacuum would give a *smaller* expectation value for the components of the stress-energy tensor, than any more realistic state in which only *parts* of space are devoid of ordinary matter. There is also good reason to suppose that in a much larger class of states, namely those which are cyclic for the local field algebras, the expectation value of the local stress-energy tensor cannot be zero (we shall consider an argument for this shortly). And the physical picture of the vacuum as a fluctuating field of zero-point energy is an extremely general one that surely does not depend on the presence of exact symmetries.

For these reasons, to object that the argument as given cannot be made mathematically rigorous appears unduly pedantic. The zero-point energy is present in elementary quantum mechanics, and it can be used to explain a large number of phenomena, from Unruh radiation to the Casimir effect; it is not going to go away because the vacuum can only be defined unambiguously in the presence of global symmetries.

The zero-point energy is therefore widely assumed to be real, and to have a very large value; and therefore that it must be cancelled (or almost cancelled) by some other physical mechanism. But the difficulties here are severe. It is a conspiracy that is needed. All of the fields in the standard model will contribute to it, so in terms of the Planck length, the cancellation will have to be fine-tuned across all these fields. It is true that fermion fields contribute with opposite sign to boson fields; were every fermion field accompanied by a bosonic partner, the cancellation would be exact. But supersymmetry, if it is a symmetry at all, is a broken symmetry. The lower bound on the mass differences of fermions and their supersymmetric partners is of the order of few hundred GeV, so whilst it is true that for energy scales much larger than this, we would expect to find cancellation of the zero-point energies, we will still have a vacuum energy with cut-off $\Lambda \sim 100$ GeV, contributing to the effective cosmological constant a term

$\lambda \sim (100\ \text{GeV})^4$—much smaller than before, but still more than 50 orders of magnitude greater than that observed.

Only one class of solutions deals with this conspiracy quite simply, namely those that are based on the anthropic principle. On such an approach one might even hope to find a conservative solution to the difficulty, using physics as is. But I believe that only the *weak* anthropic principle has any credibility: the self-evident principle that, given a range of actual environments, that include some that are hospitable to life, then it is in one of the latter that we will find ourselves situated (that our environment is hospitable to us no longer needs an explanation). Understood in this way, the principle can only be used to solve the cosmological constant problem if there exist many actual universes, or parts of one actual universe, in which the cosmological constant takes on different values, including very small ones (granted that a very small cosmological constant is necessary if there is to be life).[1] So here too one is driven to new physical principles, to underwrite this proliferation in values of λ

In this situation it is worth taking a more leisurely and critical look at the arguments for the reality of the zero-point energy. Along the way it will be helpful to gain some historical perspective on the problem; this is not the first time that the vacuum state has been thought to be filled with activity. I shall begin with two previous pictures of the vacuum which definitely do seem to be mistaken.

16.3 The Classical Ether

The history of the classical ether is familiar, so here I shall be brief. The classical models of ether gave rise to severe problems. They appeared to need contradictory mechanical properties, and, seemingly independent of those, they led to the prediction of observable effects that were not in fact

[1]It is essential that there be an *actual* plurality. Weinberg appeals to what he calls "the weak anthropic principle", but according to him in this "one explains which of the various possible eras of the universe we inhabit, by calculating which eras or parts of the universe we could inhabit" (Weinberg 1989 p.7), leaving it ambiguous whether these different eras are merely possible or whether they all actually exist. In particular he appears to allow, both here and in his most recent use of the principle, that it is sufficient to argue that the state be written as a superposition of components in each of which the cosmological constant has a different value. This would permit the use of the weak principle, as I understand it, only on the Everett or many-worlds interpretation.

detected: if not ether drag, then a positive result for the Michelson-Morely experiment.

Why was the ether thought to exist at all? A common answer is that since light and electromagnetic forces propagate as waves, it was thought that there must be a substratum which is in motion; something had to wave. But this answer is unedifying. It does not connect with the question. The point is to explain the belief in a *mechanical* ether, not in something or other which waves. (The electromagnetic field is something or other which waves.)

A better response is that in the early days of the ether, Newtonian mechanical principles were the only ones available. From the beginning, wave equations were formulated as applications of mechanical principles to a mechanical medium with movable parts. And it *remained* a problem, as the theory of the luminiferous ether progressed, to interpret and explain the equations in Newtonian mechanical terms. That is why, in extending this theory to electromagnetic phenomena, Maxwell made use of mechanical models, and why so much of late nineteenth century work on the structure of electromagnetic media was directed to the investigation of *material* dielectrics.

In Lorentz's work the treatment of material dielectrics was initially continuous with his treatment of ether, but the ether was progressively shorn of its mechanical properties. Its principal role, at the end, was to define the resting frame, to which all the electrodynamical equations were referred—and in terms of which the properties of moving dielectrics were analyzed. It is common to view the disappearance of the ether theory, following Einstein's intervention, as *abrupt*, but it would be more accurate to say that the ether was progressively whittled away. In Lorentz's hands it remained little more than a frame of reference, and a dwindling array of methods for studying differential equations in terms of difference equations. It was the grin of the Cheshire Cat.

In fact, changed into something so much more abstract, Lorentz's point of view has its defenders even today. As argued by Bell (1989), one can take the view that contraction and dilation effects should be thought of as perfectly real, when referred to a fixed inertial frame throughout, at least as a pedagogical device; that they are most clearly understood as *dynamical* effects; that they should be analyzed in terms of objective changes in the electromagnetic structure of measuring rods and clocks, referred to the one inertial frame throughout. Lorentz would have been perfectly at home with all this.

So when, precisely, did the ether disappear? Perhaps it was when McCullogh's ether, providing as it did constitutive equations for a continuum mechanical system, giving the right boundary conditions for Fresnel's ether theory, was not adopted; or perhaps it was when Maxwell, having derived the equations for the displacement current from his system of cogs and wheels, derived them instead (in the *Treatise*) abstractly, from a Lagrange function; or perhaps it was when Lorentz, struggling to derive the Fresnel ether drag coefficient from his study of moving dielectrics, exempted his molecules of ether from Newton's third law. Einstein's 1905 intervention, which of course limited Newton's force laws much more comprehensively, also showed that no privilege could attach to any *particular* inertial frame: whatever the virtues of a frame-dependent analysis of the dynamical behaviour of rods and clocks, it could not make a difference as to *which* frame was to be used. Einstein surely delivered the death blow to the hypothesis of ether, but it was the *coup de grâce*, not the *coup mortel.*

16.4 The Dirac Negative Energy Sea

Our second example is much closer to home. It is the Dirac negative energy sea. The negative energy density of this sea is, indeed, the same as that of the zero-point energy in QFT.

Dirac introduced this vacuum in order to solve the negative-energy difficulty. That in turn had plagued every previous attempt to unify quantum mechanics and special relativity, beginning with the relativistic scalar wave equation introduced by Schrödinger, Gordon, and Klein in 1926. Dirac was dissatisfied with this equation. It admitted negative-energy solutions, but further, as a second-order equation, it appeared to him inconsistent with the basic structure of quantum mechanics. In particular the charge-current equation yielded a conserved quantity that was not positive definite, so one could not use it to define a positive-definite norm. Dirac concluded, rightly, that it could not itself be interpreted in terms of probability. In 1928 he found the Lorentz covariant first-order equation that bears his name, for spin-half particles, which did yield a positive-definite norm. With that he was in a position to define a Hilbert space. But the equation still admitted negative-energy solutions, and it was clear that the Klein paradox could be formulated here as in the scalar case. In general one could not simply exclude the negative-energy solutions. The obvious candidates for interaction terms also led to transitions from positive to negative energy states.

Dirac's remarkable solution of 1930 was to suppose that in the vacuum state all the negative energy states were already filled. It then followed, from the Pauli exclusion principle, that electrons could not make any transitions from positive to negative energies. The exception, of course, was if some of the negative energy states were *not* filled. And this could come about easily enough, if only a negative energy electron is given sufficient positive energy. It would then appear as a positive energy particle, leaving behind it a *hole* in the negative-energy sea. And this hole, as Dirac argued in 1931, would behave just like an electron of positive energy but with opposite charge. In this way the concept of *antimatter* was born.

With it, automatically, came the concept of pair annihilation and creation: a negative-energy electron, given sufficient energy, appears as a positive-energy electron, leaving behind it a hole, i.e. a positron: a particle and its antimatter partner both come into being—*pair creation*. Equally, the hole can subsequently be filled: not only the hole, but the positive-energy electron that fills it disappear—*pair annihilation*. All will be in order so long as only energies *relative* to the negative energy sea have any physical significance.

These ideas translate readily into a reformulation of the canonical second quantization formalism, that Dirac had developed to treat the many-body problem in NRQM, and to treat (non-relativistically) the radiation field as a boson ensemble, some three years previously. In this formal framework, many-body operators $d\Gamma(X)$ are defined, for any 1-particle operator X, by replacing what in NRQM would be the expectation values of X by the corresponding expression in which the state (ket) ψ is replaced by a q-number field Ψ (the annihilation field), and the conjugate (bra) replaced by the adjoint field (the creation field). Thus, in the case of the Hamiltonian H, using the position representation:

$$\langle H \rangle = \int \psi^*(x,t)\ H\ \psi(x,t) d^3x \ \longrightarrow\ d\Gamma(H) = \int \Psi^*(x,t)\ H\ \Psi(x,t) d^3x. \tag{16.3}$$

Note that the RHS is a q-number, whereas the LHS is a c-number (hence *second* quantization). Note further that the annihilation field Ψ always stands to the right, so such expressions always annihilate the vacuum; the vacuum is always the zero eigenstate of any canonically second quantized quantity. And note that this correspondence can be used to define *local* field densities as well, replacing H by the Dirac delta function (not a self-adjoint operator, admittedly, but a bilinear form), again giving quantities which automatically annihilate the vacuum.

Even more so than in elementary quantum mechanics, the momentum representation has a special status. Let $b_r^*(\vec{p})$, $b_r(\vec{p})$ be, respectively, the creation and annihilation fields for a positive-energy solution to the free Dirac equation, with 3-momentum $\vec{p}$ and bispinor w_r, $r = 1, 2$. For each p the set of such bispinors is a 2-dimensional vector space, so any such solution can be written as a linear combination of these creation operators applied to the vacuum. Let $b_{r+2}^*(\vec{p})$, $b_{r+2}(\vec{p})$ be the corresponding operators for the negative energy bispinor w_{r+2}, $r = 1, 2$. Let $p_0 = +\sqrt{\vec{p}^2 + m^2}$. Then the annihilation field $\Psi(x,t)$ has the Fourier decomposition:

$$\begin{aligned}\Psi(x,t) \quad = \quad & \int \sum_{r=1,2} \Big(w_r(\vec{p}) b_r(\vec{p}) \; e^{-i(p_o t - \vec{p}\vec{x})/\hbar} + \\ & + w_{r+2}(\vec{p}) b_{r+2}(\vec{p}) \; e^{i(p_o t + \vec{p}\vec{x})/\hbar} \Big) \frac{d^3 p}{p_0}. \end{aligned} \tag{16.4}$$

It involves annihilation operators only. The Fourier expansion for the adjoint field only involves creation operators. We now follow the recipe of second quantization, (16.3), using (16.4) and its adjoint, for the one-particle Hamiltonian $H = \pm p_0$ (in the momentum representation) and for the charge operator $-e$ (a multiple of the identity). The second quantized energy and charge are, respectively:

$$d\Gamma(H) = \int \sum_{r=1,2} p_0 \Big(b_r^*(\vec{p}) \; b_r(\vec{p}) - b_{r+2}^*(-\vec{p}) \; b_{r+2}(-\vec{p}) \Big) \; \frac{d^3 p}{p_0} \tag{16.5}$$

$$d\Gamma(-e) = -e \int \sum_{r=1,2} \Big(b_r^*(\vec{p}) \; b_r(\vec{p}) + b_{r+2}^*(-\vec{p}) \; b_{r+2}(-\vec{p}) \Big) \; \frac{d^3 p}{p_0}. \tag{16.6}$$

Necessarily, in accordance with (16.3), these operators still annihilate the vacuum, since the annihilation operator automatically appears on the right.

So far everything has been done in exact correspondence to NRQM. Now for the change in the physical interpretation, and corresponding to that, a change of notation. We are going to consider that all the negative energy states are filled. Since the absence of a negative energy electron with bispinor w_{r+2} and 4-momentum $(-p_0, \vec{p})$ behaves just like the presence of a positive energy particle of opposite charge, the annihilation of the former is equivalent to the creation of the latter, and *vice versa.* Evidently the operators (16.4), (16.5) and (16.6) do not annihilate this new vacuum state. To reflect these facts, denote $b_3(-\vec{p})$ by $d_1^*(\vec{p})$, and $b_4^*(-\vec{p})$ by $d_2(\vec{p})$. Also introduce new notation for the bispinors, denoting w_r by u_r and w_{r+2} by v_r. In terms of these notational changes, the Fourier expansion for the field Ψ (before a pure annihilation field, as given by (16.4)) becomes:

$$\Psi(x,t) = \int \sum_{r=1,2} \left(u_r(p)\; b_r(p)\; e^{-ipx/\hbar} + v_r(p)\; d_r^*(p)\; e^{ipx/\hbar}\right) \frac{d^3p}{p_0}. \quad (16.7)$$

The new notation reflects the action of the field on the new vacuum. "Effectively", it is a sum of (hole) creation and (electron) annihilation fields. Concerning (16.5), (16.6), here we still want to end up with operators which annihilate the new vacuum, so let us re-order terms in these expressions so that the effective annihilation operators always stand to the right. These operators must obey *anticommutation* relations, so as to preserve the antisymmetrization of the states they act on, so this introduces a change in sign. It also introduces the c-number value of the anticommutators, which we must integrate over (a divergent integral). In the new notation, we thus obtain:

$$d\Gamma(H) = \int \sum_{r=1,2} p_0 \left(b_r^*(\vec{p})\; b_r(\vec{p}) + d_r^*(\vec{p})\; d_r(\vec{p})\right) \frac{d^3p}{p_0} - \text{ infinite constant} \quad (16.8)$$

$$d\Gamma(-e) = -e \int \sum_{r=1,2} \left(b_r^*(\vec{p})\; b_r(\vec{p}) - d_r^*(\vec{p})\; d_r(\vec{p})\right) \frac{d^3p}{p_0} - \text{ infinite constant.} \quad (16.9)$$

The infinite constants are readily interpreted as the energy and charge (both negative) of the Dirac vacuum. The change in sign makes the q-number part of the total energy non-negative, that of the total charge indefinite. Each involves number operators for the electrons out of the sea, and the holes, the positrons, in the sea. In both cases they have only positive energies.

Evidently this theory brings with it a problem exactly like the zero-point energy difficulty—a vacuum energy twice the value of the latter for each p, r, but only for the negative-energy states (so with the same total negative energy). The cancellation of the zero-point energy for fermion fields (negative) and boson fields (positive), given unbroken supersymmetry, follows from the hole theory just as from standard fermion QFT. The negative-energy fermion sea cancels the zero-point energy of the associated boson and antiboson fields.

A comparison with the classical ether is also instructive. The Dirac vacuum was formulated in terms of relatively *traditional* mechanical principles (in that interactions were introduced using phase space methods defined only for spaces of constant, finite dimensionality). As such, particle number was necessarily conserved (pair creation and annihilation processes always involved transitions between states of the same number of particles). This was hardly based on a metaphysical principle, on a par with the principle that there must be a bearer of the motion of waves (but then neither was the latter a very plausible basis for the commitment of classical ether theorists to the ether); rather, its roots, like the roots of the ether, were pragmatic: no other method was known for introducing particle interactions. It was the same in the case of Dirac's quantization of the electromagnetic field three years previously: there too, although it was clear that photon number should be subject to change, Dirac modelled such processes in terms of transitions between states which preserved particle number. In this case the transitions were to and from a sea of zero-energy photons. (Here too photon number was preserved.)

The negative energy sea was effective in other ways as well. Like the classical ether, it was a fertile source of heuristics. Dirac was quickly led to the concepts of vacuum polarization, and of contributions from the sea to the effective charge and mass of the electrons and holes. But equally, and again in parallel to the classical ether, the new vacuum did not really make physical sense. It was hard to take the theory as literally true (it was "learned trash", according to Heisenberg).

The field theory which replaced the hole theory was introduced by several authors, by Fock in 1933, by Furry and Oppenheimer in 1934, and by Heisenberg in 1934. There was no canonical second quantization. There was to be a field (and an adjoint field) now taken to be fundamental, obeying anticommutation relations which were understood as *quantization* rules. Each was written down as before as a Fourier expansion in normal modes, but now the coefficients of these expansions were interpreted *ab initio* as a combination of antiparticle creation and particle annihilation fields, exactly as in (16.7). The global operators for the field could be obtained in formally the same way as in the second-quantized theory (16.3), but the c-number expressions were understood purely classically.

It was essential that the q-number fields had the very particular action on the vacuum—no longer a negative-energy sea!—given by (16.7), acting on the Dirac vacuum. For this no explanation was given. The re-ordering process was still to be used, because with the field (16.7), entering into expressions of the form (16.3), one obtains creation fields to the right; but now the c-number values of the anticommutators were simply discarded. This process was called *normal-ordering.* It was viewed as part of the increasingly elaborate procedure for isolating finite expressions in perturbation theory. Only one infinite negative term was allowed an occasional physical explanation: a (negative) zero-point energy.

More common was to reserve for physical interpretation only the normal-ordered quantities. In terms of the normal-ordered energy, the vacuum of the field theory has zero energy. For generations now physicists have shifted back and forth between the view that the subtractions of renormalization theory are no more than formal, and the view that they reflect real physical quantities. The Dirac vacuum provided a clear physical picture of *all* these subtractions, based at it was on the canonical formalism of NRQM, and, apart from certain notable exceptions, at least for one generation of physicists—Dirac's generation—this picture had become the fundamental one. Witness Wightman:

> It is difficult for one who, like me, learned quantum electrodynamics in the mid 1940s to assess fairly the impact of Dirac's proposal. I have the impression that many in the profession were thunderstruck at the audacity of his ideas. This impression was received partly from listening to the old-timers talking about quantum-electrodynamics a decade-and-a-half after the creation of hole theory; they still seemed shell-shocked. (Wightman, 1972 p.99)

One might add that Wightman never had to *accept* the hole theory; he never had to work with it. Fiction is never shell-shocking.

A further blow for the Dirac vacuum came with Pauli and Weiskopf's treatment of the complex *scalar* field, using *commutator* relations, in 1934—with a similar interpretation of the Fourier expansion as the electron-positron field, with similar normal ordering prescriptions, and with its application to the new field of meson physics. With that antiparticles were seen as ubiquitous; they were the appropriate field-theoretic account of the negative energy terms, likewise ubiquitous in relativistic quantum theory, with no special connection to the Pauli exclusion principle. Obviously with scalar fields there could be no question of a filled negative energy sea.

There is a last chapter to this story, but a more contentious one. To compare it once more to the history of ether, the standard post-hole theory view of the vacuum—as presented, say, in Wentzel's influential textbook introduction to the subject—was much closer to Lorentz's view of the ether than to Einstein's. (Quantum) mechanical principles are no longer applied to the QED vacuum; the explanations that it offers of pair creation and annihilation events, in NRQM terms, are no longer taken literally; but still there lurks a story of sorts to be told, to account for the plane-wave expansion (16.7), just as there lurked a story to be told to explain the length contraction and time-dilation effects in classical electromagnetism. There is an analog to the Lorentz pedagogy; call it the *Dirac pedagogy.* Even very recent introductions to the subject make use of it. The Dirac vacuum is not taken realistically—indeed, it is sometimes introduced without even mentioning the hole theory—no more than the Lorentz pedagogy takes seriously the resting frame, or makes mention of ether. But what is missing is an *alternative* account of the plane wave expansion, and of the details of the relationship of negative energy states to antiparticles, and of the meaning of normal ordering. There is as yet no good analog to the Einstein-Minkowski *geometric* account of contraction and dilation phenomena, in terms of invariant intervals between events (that they are the consequences of taking different events as the simultaneous end-points of rods, and different events as the simultaneous ticks of clocks, depending on one's choice of simultaneity).

Is there an alternative explanation for the plane-wave expansion—for how antiparticles get into the theory? It can certainly be shown that only fields built up out of creation and annihilation fields for two kinds of particles, as given by (16.7) and its adjoint, can be Lorentz-covariant, satisfy microcausality, and transform simply under a U(1) gauge symmetry (Weinberg 1964, Novozhilov 1975). The two kinds of particles have to be identical in all respects, save that they have opposite charge. On this approach one starts from the free one-particle Hilbert space theory, using the Wigner classification of the irreducible representations of the Poincaré group. Creation and annihilation operators can be defined in these terms over the associated Fock spaces, just as in NRQM. But one never in this way makes any mention of negative-energy states, and the normal ordering process is unexplained.

Weinberg's account of the structure of free-field theory is on the right lines, but it can certainly be improved on. For this we need Segal's method of quantization, itself a fragment of what is nowadays called geometric quantization. Given a complex structure J on the classical solution manifold V of a linear system of equations—a linear map such that J^2 acts as minus the identity—and given a non-degenerate bilinear form S on V, one can always define a Hilbert space V_J, and from this construct a Fock space $F(V_J)$ over V_J. If these equations and the bilinear form are covariant, this Hilbert space will inherit their covariance group. In the symmetric case, define the *Segal field* abstractly, as a linear map Φ from V to self-adjoint operators on Hilbert space, obeying the anticommutator:

$$\Big\{\Phi(f), \Phi(g)\Big\} = \hbar S(f, g). \tag{16.10}$$

A field with these properties can be represented concretely, given creation and annihilation fields on $F(V_J)$, by the relations:

$$\Psi_J(f) = \frac{1}{\sqrt{2\hbar}}\Big(\Phi(f) + i\Phi(Jf)\Big) \tag{16.11}$$

$$\Psi_J^*(f) = \frac{1}{\sqrt{2\hbar}}\Big(\Phi(f) - i\Phi(Jf)\Big). \tag{16.12}$$

From the anticommutator (16.10), Ψ_J and Ψ_J^* obey the anticommutators characteristic of annihilation and creation operators on the antisymmetric Fock space, and *vice versa.* These anticommutators vanish for f, g with spacelike separated supports, if S does.

If one begins with Poincaré-covariant complex equations, there are always two possible choices of J. One of them is just multiplication by i, denote J_N. It is local, but the energy it gives rise to (when used in Stone's theorem, for the generator of translations in time) is indefinite. The other choice, denote J_P, makes use of the decomposition of the classical solutions into positive and negative frequency parts (positive and negative "energies"). This is non-local, but gives a positive energy. The two are related by the Segal field Φ, which is independent of the complex structure. One then finds that (with $f^+(f^-)$ the positive (negative) frequency part of f):

$$\Psi_{J_N}(f) = \Psi_{J_P}(f^+) + \Psi^*_{J_P}(f^-) \tag{16.13}$$

$$\Psi^*_{J_N}(f) = \Psi_{J_P}(f^-) + \Psi^*_{J_P}(f^+). \tag{16.14}$$

The quantities on the left, defined with respect to the local complex structure J_N, are the local, causal fields; they have exactly the interpretation we have been seeking, in terms of the creation and annihilation operators defined using the non-local complex structure. (16.13) is the abstract analogue of (16.7). Dirac was canonically second-quantizing using the local complex structure, so using the Ψ_{J_N}'s in (16.3). By linearity, this ensures the invariance of such terms under rotations in the local complex structure. This in turn forces charge conservation. But he used the same complex structure to define their particle interpretation, their Fock-space action. In this way he forced number conservation too, but at the price of introducing negative energies. These same interaction terms are not invariant under rotations in the non-local complex structure, the one which ought to be used to define the particle interpretation. Using the latter, particle number is not preserved.

Introducing the negative energy sea is in fact a way to switch between one complex structure and the other. Further, if one normally-orders canonically second quantized operators, defined on $F(V_{J_N})$, one obtains the canonically second quantized operators, defined on $F(V_{J_P})$, so the normal ordering process likewise switches between the two complex structures (Saunders 1991). But now it follows that the canonical second quantization cannot be used to define local bilinear quantities, in particular couplings to other fields, so this correspondence is not of much use outside of free-field theory.

This same framework applies to boson fields, save that there one has an *antisymmetric* bilinear form (the symplectic form). The Segal field in this case satisfies commutation relations, but otherwise the same analysis goes through; normal ordering again switches between one complex structure and the other. There is only this difference: one cannot switch between them by modifying the vacuum, and one can no longer interpret normal ordering in terms of a filled negative energy sea (for of course the exclusion principle no longer applies). The fact that the non-local complex structure is the one that is used at the level of the Hilbert-space theory, in both cases, also explains why there is no local, covariant position operator in relativistic quantum theory (Saunders 1992). The analysis applies equally to NRQM, save that there the two complex structures coincide.

With that the Dirac vacuum can, I think, finally be laid to rest. We have not replaced it with a better physical picture, but we have shown why any linear, Lorentz covariant system of equations, used to define a one-particle Hilbert space and a Fock space over that, will lead to the two systems of fields, related in a way which can be explained using the hole theory (in the case of a symmetric bilinear form). We see why, both for bosons and fermions, the negative energy states are associated with antiparticles; we see why interactions built up from local, covariant, gauge invariant quantities, automatically lead to pair creation and annihilation processes. And we have some insight into the meaning of the normal ordering process.

I do not believe that there is a negative energy sea. Does it follow that there is no energy in the vacuum? Can we not, as in non-relativistic quantum theory, suppose that really we have a many-particle theory, that all the physically meaningful quantities can be written as quantities of the form $d\Gamma(X)_{J_P}$, zero in the vacuum of $F(V_{J_P})$? But we have already noted that the local quantities cannot be obtained in this way. And, of course, bilinear quantities constructed from the local fields (16.13), (16.14) have as before formally divergent expectation values in the vacuum. Such quantities cannot be normal-ordered (normal ordering is a global procedure). We have transformed the problem of vacuum, we have not made it go away.

16.5 Zero-Point Energy

If one takes quantum fields as fundamental, rather than as devices for introducing local particle interactions, then there is a clear intuitive argument for supposing the ground state has non-zero energy, deriving from the elementary theory of the harmonic oscillator.

This is again non-local, depending as it does on the Fourier decomposition of the field into normal modes, and with that the particle interpretation. A more direct argument is possible, which shows why the components of the local stress-energy tensor cannot have zero-expectation value in the Minkowski space vacuum. Certainly no quadratic expressions in local self-adjoint quantities, and no quantities of the form $\Psi^*\Psi$, bilinear in the local fields, can yield zero in this state: all such expectation values must be strictly positive, as a corollary of the Reeh-Schlieder theorem.

To see why no finite subtraction can help, consider the CCR's for the components of the stress-energy tensor $T^{\mu\nu}$:

$$\left[T^{0k}(x), T^{00}(x')\right] = -iT^{00}(x)\,\partial^k\,\delta(\mathbf{x}-\mathbf{x}') - iT^{kl}(x')\,\partial_l\,\delta(\mathbf{x}-\mathbf{x}') \quad (16.15)$$

(see Schwinger 1973 p.26). If we replace T by $\widetilde{T} = T + \lambda R$, for some matrix of real numbers R, then from the commutator it follows that R will have to be a multiple of the Minkowski metric g. But in that case, if we require that all the components of $\widetilde{T}$ vanish in the vacuum:

$$\langle \widetilde{T}^{00} \rangle = \langle \widetilde{T}^{0k} \rangle = \langle \widetilde{T}^{kl} \rangle = 0 \quad (16.16)$$

then certain constraints follow, among them:

$$\left\langle T^{00} + \frac{1}{3}T^{kk} \right\rangle = 0. \quad (16.17)$$

(using the Einstein summation convention). In the case of the electromagnetic field, $T^{kk} = T^{00}$, so:

$$\frac{4}{3}\left\langle T^{00} \right\rangle = 0. \quad (16.18)$$

Since T^{00} is $\frac{1}{2}(\vec{E}^2 + \vec{B}^2)$, a sum of terms quadratic in local self-adjoint quantities, equation (16.18) cannot be satisfied, on pain of violating the Reeh-Schlieder theorem (supposing the vacuum is a cyclic vector).

So much for pure electromagnetism; similar arguments apply to other examples; the difficulty appears quite general. One can of course always renormalize the expectation value of T by an infinite subtraction. So much is routine for most quantities of physical interest in quantum field theory, where divergent expressions are obtained by almost any naive use of the formalism. Using the more modern methods of renormalized perturbation theory, where one introduces counterterms at each order, one can even obtain finite contributions from the perturbation considered is a smooth and controllable way. But it was always clear that in non-gravitational physics, only *changes* in energy are significant. The whole philosophy of renormalization theory is built on this principle. A theory is to be defined at a certain scale M; the conventions there adopted remove all ultraviolet divergences by *fiat.* One then imagines a shift in the scale M, with a corresponding shift in the renormalized coupling constants, field strengths, and Green's functions. The renormalization group equation tells us how these shifts are interrelated. One never considers the absolute values, but just these are what are relevant, uniquely, to gravity. The question is as before: is there reason to think that there is an absolute energy density associated with the vacuum, of an unreasonably large nature?

The Casimir effect is often invoked to show that there is. It is worth looking at this effect in some detail. Recall that Casimir predicted, in a now celebrated paper published in 1948, that an attractive force should act between two uncharged parallel conducting plates in vacuuo. If their separation is L, the force per unit area is:

$$P = \hbar c \, \frac{\pi^2}{240} \, \frac{1}{L^4}. \tag{16.19}$$

In deriving this result, Casimir computed expressions for the zero-point energy, evaluated with and without (Dirichlet) boundary conditions for the fields on the wall, using a cut-off to control the high frequency behavior. The difference ΔE is a function of the separation L of the walls; on variation of L one obtains the pressure. In fact, since Casimir reasoned that wavelengths $\lambda > L$ are excluded, as L increases the energy should increase. The force that results is therefore attractive.

This was in itself not at all surprising: the existence of attractive electrostatic forces between neutral bodies was familiar (Van der Waals forces), and indeed it was with this idea in mind that Casimir was led to his discovery. He began by trying to calculate the Van der Waal's force between two poloariable atoms. To simplify the analysis, he replaced one of the atoms by a conducting plate; from there he was led to consider the even simpler problem of two parallel plates. Casimir found, to get agreement with experiment, that the influence of these charge polarization effects must be treated as propogating at the speed of light, suggesting the existence of an energy density in the space between the plates. He then found that the long range limit of the force can be derived entirely by considering the change in this field energy.

Shortly after Casimir's pioneering work, there appeared a second quantitative account of the effect. The model was due to Lifschitz. He suppposed there to be fluctuations in the polarization fields associated with the electrons in the conducting plates (realistically, no metal is a perfect conductor at arbitrarily high frequencies), and that these should couple with each other, giving rise to van der Waals forces. The coupling can be modelled either in terms of retarded distance forces, or as mediated by electromagnetic fields, with these fluctuating polarization fields as sources.

Explicitly, Lifshitz introduced a random polarization field for a material of dielectric constant ϵ_0; the expression he derived for the resulting pressure, from their mutual attraction, is:

$$P = \frac{\hbar c \pi^2}{240 a^4} \left(\frac{\epsilon_0 - 1}{\epsilon_0 + 1} \right)^2 \varphi(\epsilon_0) \tag{16.20}$$

where φ is a function with the limiting behavior, for $\epsilon_0 \rightarrow \infty$:

$$\varphi(x) \rightarrow 1 - (\frac{1.11}{\sqrt{\epsilon_0}}) \ln(\frac{\epsilon_0}{7.6}). \tag{16.21}$$

In this limit (16.20) and (16.19) coincide. Admittedly this limit cannot be taken literally, as in the limit of a perfect conductor there will be no polarization field, but no more did Casimir's method make sense in this limiting case. He only obtained a finite value of the vacuum energy, in the presence of the plates, by introducing a a high-frequency cut-off to off-set their influence.

Indeed, the two explanations can be viewed as perfectly compatible. The quantum vacuum can be interpreted as just such a system of fluctuating fields, of the sort invoked by Lifschitz. Whence the random fluctuations in the dielectric medium, in the zero-temperature limit? On general physical principles, such fluctuations should be dissipative; if so, equilibrium can only be reached if there are likewise fluctuations in the electromagnetic field in the cavity. This fits well enough with Casimir's interpretation, according to which, even at zero temperature, the vacuum has non-zero energy.

Lifshitz's argument can thus be seen to strengthen Casimir's. In subsequent years, the discovery of the extensive formal parallels between quantum field theory and condensed matter physics has only strengthened the case: where in condensed matter physics one has thermal fluctuations, in quantum field theory one has vacuum fluctuations. Indeed, a Casimir-like force has been shown to act between parallel plates separated by matter at a bulk critical point (Fisher and de Gennes 1978; see also Cardy 1990). At the critical point the correlation length becomes large, so long-range thermal fluctuations are set up in the medium; these correspond to the vacuum correlation functions in quantum field theory. However, unlike the latter case, here there is no problem with divergences (for in condensed matter physics one has a natural physical cut-off). Nor is there any doubt that the energy of the thermal fluctuation should gravitate.

But now a problem arises for both interpretations: the Casimir effect does not always give rise to an attractive force. Later calculations showed it to be *repulsive* in the case of spherical shells, a discovery which killed off a speculative attempt to calculate the fine-structure constant $e^2/\hbar c$ from first principles. The old Abraham-Lorentz model, recall, supposed the electron, in its rest frame, to be a charged conducting spherical shell of radius a. Its electrostatic energy is

$$E_e = e^2/2a \tag{16.22}$$

with a corresponding tension $e^2/8\pi a^4$, tending to expand the shell. In view of his results for parallel plates, Casimir suggested that there would be an energy E_c associated with the sphere, which would increase with its radius, so giving rise to a compensating attractive force. On dimensional grounds this will be of the form:

$$E_c = -C\ (\hbar c/2a). \tag{16.23}$$

The resultant tension in the surface will vanish if C equals the fine structure constant. Using the parallel plates result, approximating a sphere of radius a as two plates of area πa^2 a distance a apart, one finds

$$C \approx 0.09 \tag{16.24}$$

which is only a factor 10 off from the fine-structure constant.

This rough calculation is numerically correct; C *is* of the order 0.09 (Boyer 1968). But the energy *decreases* with increasing radius; the Casimir force is repulsive, in the case of a sphere, not attractive. More generally, the Casimir effect turns out to be extremely sensitive to the geometry of the boundaries involved, as well as their composition (and to the space-time dimension, curvature, and type of field).

The fact that repulsive forces can be obtained has been cited as a reason for rejecting the interpretation of the effect in terms of van der Waals forces (Elizalde and Romeo 1991), on the grounds that the latter are always attractive. This claim has been disputed, however; it does not appear to apply to the Lifschitz model. And such forces equally pose a difficulty for Casimir's interpretation of the effect: how can they be repulsive, if the presence of the conducting walls *excludes* normal modes of the field?. In itself, the dependence of sign on geometry does not settle the question of what explanation is to be preferred, although it considerably increases theoretical interest in the effect.

I have said that Lifschitz's approach can be reconciled with Casimir's; can it also be viewed as a genuine alternative? It was understood in this way by Julian Schwinger. By the mid '60s, Schwinger, dissatisfied with the mathematical foundations of renormalization theory, had recast a number of important calculations in quantum field theory in terms of classical field theory. In this, the so-called "source" theory, the vacuum state was to be viewed as having strictly zero energy. Schwinger first treated the Casimir effect in these terms in 1975; he and his co-workers presented more refined studies in 1977 and 1978. It was granted that "the Casimir effect poses a challenge for source theory, where the vacuum is regard as truly a state with all physical properties equal to zero". Here is the approach they recommended:

> The Casimir effect is a manifestation of van der Waals forces under macroscopic circumstances. When material, uncharged bodies are present, whether dielectrics or conductors, nonvanishing fluctuating electric and magnetic fields are generated, which give rise to stresses on the objects. These fields may be most easily expressed in terms of Green's functions evaluated in a source theory formulation of quantum electrodynamics (Milton et al. 1977).

To give a flavour of the analysis, one starts from the electric and magnetic fields defined by the source field $\vec{P}$, just as for a polarization field:

$$\nabla \times \vec{H} = \epsilon \dot{\vec{E}} + \dot{\vec{P}}, \quad \nabla \cdot (\epsilon \vec{E} + \vec{P}) = 0. \tag{16.25}$$

The fundamental objects in the theory are the Green's functions $\vec{\Gamma}(x, x')$ by means of which the fields and sources are related:

$$\vec{E}(x) = \int \vec{\Gamma}(x, x') \cdot \vec{P}(x') d^3 x^{prime}. \tag{16.26}$$

The general method is to consider the change in the action (or the energy), expressed as integrals over the sources, on variation of the parameters determining the geometry of the dielectrics: by subtracting the vacuum Green's function (spherical case), and by variation of the dielectric constant (parallel conductors, corresponding to a change in the distance between the plates). In the latter case Schwinger discards a term which he identifies as the change in vacuum energy. The change in energy defined by the volume integral of the fields (including the polarization fields), due to variation of the dielectric constant $\delta\epsilon$ is:

$$\delta E = \frac{i}{2} \int \delta\epsilon(x, \omega)\ \Gamma_{zz}(x, x, \omega) \frac{d\omega}{2\pi} (d^3 x). \tag{16.27}$$

The polarization field $\vec{P}$ has dropped out of the analysis. What remains is the computation of the Green's function; the component Γ_{zz} is given by the expression:

$$\Gamma_{zz} = \left[\ \omega^2 g^E + \frac{k^2}{\epsilon\epsilon'} g^H + \frac{1}{\epsilon} \frac{\partial}{\partial z} \frac{1}{\epsilon'} \frac{\partial}{\partial z'}\ g^H \right] |_{z=z'} \tag{16.28}$$

where the g's are the Green functions for the electric and magnetic field satisfying, for the electric field:

$$\left[-\frac{\partial^2}{\partial z^2} + k^2 - \omega^2 \epsilon \right]\ g^E(z, z') = \delta(z - z') \tag{16.29}$$

(the equation for the magnetic case is similar). Of the term contributing to δE, which Schwinger interprets as the change in volume energy, he says: "Since this term in the energy is already phenomenologically described, it must be cancelled by an appropriate contact term". But it is not clear that this is correct; the analysis throughout has been in terms of the phenomenological quantities; the maneuver appears to be *ad hoc.*

The remaining method used by Schwinger, whereby the radial component of the stress-energy tensor for the electromagnetic field is calculated, likewise involves an infinite subtraction. It is justified with the words: "No physical meaning can be ascribed to such a term, however, since no stress can arise from a homogeneous dielectric (as such it can be canceled by a contact term)". This term too has the same form as the expression for the vacuum energy arising in Casimir's calculation.

Whatever the virtues of Schwinger's source theory, transparency, and statements of clear and systematic principles, are not among them. I do not believe his methods deliver an unambiguous verdict on this matter.

The essential question remains: are the conductors the source of the vacuum energy, if any, in the region bounded by the conductors? If so there is no evidence, coming from this quarter, for the zero-point fluctuations. The argument that it is not is as stated: no equilibrium would then seem to be possible. By the fluctuation-dissipation theorem, the system will be dissipative. The principle is an extremely general one.

Against this Rugh, Zinkernagel and Cao (1999), in their review of the various treatments of the Casimir effect, have suggested:

> The answer seems to be that there is no place for the dissipated energy (from the dipole) to go. If there is no vacuum field (in standard QED, an ensemble of harmonic oscillators) where the energy can be dissipated into, then there is no vacuum-dipole fluctuation-dissipation theorem.

Their proposal will have to be more radical than merely repudiating the vacuum of QED, however. They will have to deny the reality of classical fields as well. This is not a claim that can be underwritten by the source theory; Schwinger did not insist that c-number fields too were unreal.

Rugh *et al.* also make another suggestion: why not adopt Lifschitz's view, accept that equilibrium is maintained by appeal to vacuum fluctuations in the space between the plates, but suppose that these fluctuations are brought into existence by the plates? But that will hardly do, unless such fluctuations are brought into existence not only between the plates, but everywhere in space; for the combined system will in that case be dissipative, with nothing beyond it to restore equilibrium.

16.6 Against Zero-Point Fluctuations

We cannot lightly abandon general arguments on the nature of thermodynamic equilibrium. Probed in the right way, the vacuum surely does offer evidence of stochastic activity. There is no doubt that the Casimir effect can be interpreted in terms of it, and it seems to me that one does have to associate an energy with the region bounded by the plates. The picture of the field as a fluctuating medium is a natural accompanient to this:

> How can a unit volume of empty space contain energy? The answer in part lies in the fact that, according to quantum mechanics, physical quantities tend to fluctuate unavoidably. Even in the apparent quiet of the vacuum state pairs of particles are continuously appearing and disappearing. Such fluctuations contribute energy to the vacuum (Abbott 1988).

On the other hand, nothing in the Casimir effect *forces* the interpretation of the vacuum energy in terms of vacuum fluctuations—nor, indeed, does the Casimir effect imply that there exists a non-zero vacuum energy outside the plates.

This is clearer in the modern treatment of the effect, where one computes the expectation value of $T^{\mu\nu}$ in the slab vacuum, imposing Dirchlet boundary conditions, and in the Minkowski space vacuum: it is the difference between these two quantities that is finite, and which is used (as a function of the dimension of the slab) to compute the Casimir force. There are, moreover, *two* distinct kinds of effect that emerge from these calculations: the one is an energy density in the neighbourhood of the boundary, which is sensitive to the precise form of $T^{\mu\nu}$ (whether for electromagnetism or a scalar field, say, and whether for the canonical or conformally coupled field); the other is uniform throughout the slab volume, which arises, from a mathematical point of view, purely through space quantization—from the fact that an integral over normal modes of the field is replaced by a sum (Fulling 1989 p.103-4). The former, if anything is, is to be traced to dynamical effects of the fields on the boundary; the latter appears to derive entirely from the fact that one has a space of finite size. One can deduce from this that there may well be a Casimir energy in the non-flat case (for closed spaces)—a conjecture borne out by more recent calculations—but there is nothing in this analysis to suggest there will be a vacuum energy in a flat space without any boundaries.

The Casimir effect shows us how the energy of the vacuum can be shifted, as a finite volume effect; it gives us no indication of its absolute value, and no reason why we should not set it to zero in flat space. "For it to be shifted it must already be real"—perhaps; but we have met a similar line of reasoning in the case of the ether ("something must wave").

There are further parallels with the case of ether. Earlier I suggested that it was not, in fact, such general intuitions ("something must wave") which drove the ether program; it was rather the felt need to understand the Fresnel (and later Maxwell) equations in terms of mechanical principles. In quantum mechanics, there is a felt need to explain the equations in terms of some underlying stochastic principles. And once the vacuum is interpreted as a fluctuating system of fields, it seems there must be an absolute energy associated with it; the existence of fluctuations is not something relative. And if this is to provide a *general* interpretation of quantum theory, if fluctuating at all, then a quantum field had better be fluctuating at all length scales, at least down to the Planck length.

The idea of quantum fluctuations certainly provides a versatile explanatory tool. We have considered it in application to the Casimir effect. It has also been used to explain the Unruh effect. Unruh showed that an accelerating detector in the Minkowski space vacuum will register a thermal background. As Sciama (1991) and others have argued, this background would be expected if the detector is sampling the zero-point spectrum along non-inertial trajectories.

The argument is semiclassical: if one evaluates the 2-point correlation function for a classical fluctuating scalar field, along a world-line of constant acceleration, one obtains a thermal distribution at the Unruh temperature. The parallel with ether is that stochastic dynamical principles now play much the same role as then did mechanical principles. A surprisingly large fragment of quantum electromagnetic phenomenology can be derived from them. By Sudarshan's theorem, one has to go to 4-point correlations to find an application that cannot be treated in semiclassical radiation theory.

All the same, my suggestion is that these semiclassical principles are not fundamental. If the precursors of ether and negative energy sea are anything to go by, we should dispense with them.

I have stated the parallel with ether. In the case of the Dirac theory, the negative energy sea is the price that one pays to preserve the familiar canonical formalism of (quantum) mechanics, in which particle number is constant. In both cases a presumed structure to the vacuum was used to extend the reach of familiar physical concepts into a novel terrain; as in quantum field theory, where a stochastic, fluctuating vacuum extends the reach of classical stochastic theories. This presumed structure to the vacuum led to internal tensions; they were resolved not by simply jettisoning the medium, but by rethinking the principles that led to it. There is plenty of evidence for statistical fluctuations in the quasiclassical limit, in thermodynamic media: the challenge is to reconcile them with a picture of vacuum in quantum field theory in which the coupling with gravity is small.

I have said there is evidence for the existence of fluctuations in the vacuum state, when probed in the right way; it is another matter to suppose they are real whether or not the experiment is conducted. Are they still there when nobody looks? One only has to put the matter in these terms for it to appear immediately in doubt: the question is too closely linked to the general interpretational problem of quantum mechanics. To invoke vacuum fluctuations in the ultra-relativistic regime, independent of any measurement context, is to take the uncritical view that ultimately the world is full of stochastic behaviour. It is the world that stochastic hidden variable theories aim to describe. But on every other of the major schools of thought on the interpretation of quantum mechanics—the Copenhagen interpretation, the pilot-wave theory, and the Everett interpretation—there is no reason to suppose that the observed properties of the vacuum, when correlations are set up between fields in vacuo and macroscopic systems, are present in the *absence* of the establishment of such correlations.

This is clear in the Copenhagen interpretation, if for no other reason than on this approach no statement at all can be made which is not tied to a definite experimental context. It is also true on the Everett interpretation, at least on those versions of it in which the preferred basis is not viewed as fundamental, in which it is significant only at the low-energy and macroscopic scale (these are the versions which attempt to relate it to the structure of the observer).[2] On this approach there are no probabilistic events underlying our macroscopic environment, unless and insofar as correlations have been established with it. And it is true on the pilot-wave theory as well—trivially, because this theory is deterministic, but also in the more substantive sense that according to it there is no dynamical activity in the vacuum at all. Take, for simplicity, the harmonic oscillator. The phase of the ground-state is independent of position, so it can be chosen as a constant. In this circumstance the particle—I am talking of the pilot-wave theory of NRQM—is stationary. Nothing at the level of the beables is in motion. It is the same with the c-number fields in the pilot-wave theory of QFT: nothing is moving in the vacuum. (This is true even though the uncertainty relations show there is statistical dispersion.) And as for the "effective" collapse of the wave-function in the pilot-wave theory (where one gets rid of terms entering into the overall superposition which no longer significantly contribute to the quantum potential), this is governed by decoherence theory just as in the Everett approach. There is no reason to extend it to the ultra-microscopic. Nothing like this need go on in the ultraviolet limit.

It follows from this that the suggestion of Rugh *et al.* may be on the right lines after all, save that it is not that there is no field in the vacuum,

[2] As such they solve the preferred basis problem by appeal to the weak anthropic principle: different patterns of correlations, corresponding to different bases, are considered as different kinds of environment. See Saunders (1993), Wallace (2002).

ready to act as a sink for the fluctuating fields established between the plates; it is that fluctuations only extend as far into the ultraviolet limit insofar as there are mechanisms in play that lead to the decoherence of these modes.

16.7 Outlook

It is one thing to modify the physical picture of the vacuum, another to fill it out quantitatively. If the stochastic background is real, but only down to lengthscales at which there is still "effective" state reduction, this will make for a contribution to the cosmological constant even in a spacetime which is spatially flat (the realistic case). How is it to be calculated? An order of magnitude argument is encouraging; one would not expect effective state reduction, outside of the laboratory context, at length scales much shorter than nuclear dimensions; indeed a cut-off at this length, of say $10^{-15}m$, yields an effective cosmological constant consistent with its observed value. But a more detailed argument is wanting.[3]

Failing a fuller analysis, the proposal is a modest one. It is that a certain physical motivation to think the zero-point energy is both real and unreasonably large, even in flat space, should be resisted. The question is whether there is any other physical ground on which to think it so.

I return to the cosmological constant problem, and specifically to the problem of how to explain the cancellation of the zero-point energy by the cosmological constant to such an extraordinary accuracy. Why should we not view this as a renormalization of the expectation value of the stress-energy tensor, a formally divergent quantity, required here just as with every other physically meaningful quantity? In the absence of a cut-off, such renormalizations must actually take place to *infinite* accuracy—what is special about the stress-energy tensor? It is not as though there is no satisfactory renormalization procedure available. The non-flat case does of course present a number of difficulties, but here there has been significant progress. Following Wald, it is possible to show that in the general case, in the absence of any symmetries, one can still define a renormalized stress-energy tensor, satisfying very natural criteria, that is essentially unique:

[3] As a further speculation, it may be that here is a basis to explain the observed approximate parity between the energy density of vacuum and matter, the so-called "cosmic coincidence problem" (see e.g. Vilenkin 2001).

"essentially" meaning that ambiguities in its definition can all be absorbed into the gravitational coupling constants. The cosmological constant, in particular, is rescaled to offset any contribution deriving from a term locally proportional to the Minkowski space metric. (See Fulling 1989, Wald 1994, for systematic expositions.)

The latter cancellation is, of course, the one usually cited in the statement of the cosmological constant problem, in the presence of a cut-off at the Planck scale. The particular difficulty of the renormalization of the stress-energy tensor is only clearly in view given this cut-off. With other quantities of physical interest, the divergences are logarithmic; the corrections they introduce remain small, even when the cut-off is taken at the Planck length. It is only when the cut-off is taken to infinity that they all become unreasonably, in fact infinitely, exact, along with that of the zero-point energy. Stopping at the Planck length, only the renormalization of the stress-energy tensor appears unreasonably exact.

Evidently our proposal answers this formulation of the difficulty as well. But the cut-off introduced to determine the gravitating energy density is significantly smaller than that used in the renormalization of other physical quantities. The implication, indeed, would appear to be that gravity is *intrinsically bound up with "effective" state reduction* (and perhaps as an equilibrium condition, as Jacobson (1995) has suggested). Evidently here, as with our historical examples, new concepts will be needed as well.

Acknowledgements

I would like to thank John Cardy and Antony Valentini for helpful discussions

References

Abbot, L. (1988): The Mystery of the Cosmological Constant, *Scientific American*, May p.82-88.

Bell, J. (1989): How to Teach Special Relativity. In *Speakable and Unspeakable in Quantum Mechanics*, Boyer, T. (1968): Quantum Electromagnetic Zero-Point Energy of a Conducting Spherical Shell and the Casimir Model for a Charged Particle, *Physical Review* **174**:1764-76.

Brandenberger, R. (1999): Back Reaction of Cosmological Perturbations and the Cosmological Constant Problem, Brown-HET-1180.

Callen, H. and T. Welton (1951): Irreversibility and Generalized Noise, *Physical Review* **83**:34-40.

Cardy, J. (1990): Universal Critical-Point Amplitudes in Parallel-Plate Geometries, *Physical Review Letters* **65**:1443-45.

Coleman, S. (1988), *Nuclear Physics B* **307**:867.

Elizalde E., and A. Romeo (1991): Essentials of the Casimir Effect and its Computation, *American Journal of Physics* **59**:711-19.

Fisher, M., and P. de Gennes (1978): *Comptes Rendus de l'Académie des Sciences*, Series B, **287**:207.

Fulling, S. (1989), *Aspects of Quantum Field Theory in Curved Space-Time*, Cambridge: Cambridge University Press.

Jacobson, T. (1995): Thermodynamics of Spacetime: The Einstein Equation of State, *Physical Review Letters* *75*: 1260-1263.

Lifschitz, E. (1956): The Theory of Molecular Attractive Forces Between Solids, *Soviet Physics JETP* **2**:73-83.

Milton, K., L. DeRaad, and J. Schwinger (1978): Casimir Self-Stress on a Perfectly Conducting Spherical Shell, *Annals of Physics* **115**:388-403.

Novozhilov, Y. (1975): *Introduction to Elementary Particle Theory*, Oxford: Pergamon Press.

Rugh, S., H. Zinkernagel, and T. Cao (1999): The Casimir Effect and the Interpretation of the Vacuum, *Studies in the History and Philosophy of Modern Physics* **30**:111-39.

Saunders, S. (1991): The Negative Energy Sea. In S. Saunders and H. Brown, eds., *Philosophy of Vacuum*, Oxford: Clarendon Press.

Saunders, S. (1992): Locality, Complex Numbers, and Relativistic Quantum Theory, *Proceedings of the Philosophy of Science Association*, Vol.1, p.365-380.

Saunders, S. (1993): Decoherence and Evolutionary Adaptation, *Physics Letters A* **184**:1-5.

Sciama, D. (1991): The Physical Significance of the Vacuum State of a Quantized Field. In S. Saunders and H. Brown, eds., *Philosophy of*

Vacuum, Oxford: Clarendon Press.

Schwinger, J. (1975): Casimir Effect in Source Theory, *Letters in Mathematical Physics* **1**:43-47.

Schwinger, J. (1973): *Particles, Sources, and Fields*, Vol.1, Reading, Mass.: Addison Wesley.

Schwinger, J., L. DeRaad, and K. Milton (1977): Casimir Effect in Dielectrics, *Annals of Physics* **115**:1-23.

Vilenkin, A. (2001): Cosmological Constant Problems and their Solutions, *arXiv:hep-th/0106083 v2.*

Wald, R. (1994): *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*, Chicago: University of Chicago Press.

Wallace, D. (2002): Everett and Structure, *Studies in the History and Philosophy of Physics*, forthcoming.

Weinberg, S. (1964): Feynman Rules for Any Spin, *Physical Review D* **133**: 1318-1332.

Weinberg, S. (1989): The Cosmological Constant Problem, *Reviews of Modern Physics* **61**:1-23.

Weinberg, S. (2000): A Priori Probability Distribution of the Cosmological Constant, *Physical Review D* **61**:103505.

Wightman, A. (1972): The Dirac Equation. In *Aspects of Quantum Theory*, Cambridge: Cambridge University Press.

This page is intentionally left blank

Chapter 17

Two Comments on the Vacuum in Algebraic Quantum Field Theory

Miklós Rédei
Eötvös University, Budapest

Abstract. *Two features of the vacuum state in algebraic quantum field theory (AQFT) are reviewed: local faithfulness (Reeh-Schlieder theorem) and the spacelike correlations it predicts. The standard interpretation of the Reeh-Schlieder Theorem, endorsed in this comment, is that it renders meaningless any talk about particles as strictly localized (in spacetime) entities. This interpretation is further supported by Malament's 1996 result (also reviewed in the paper) asserting that there exists no non-trivial, covariant position observable on any spacelike hypersurface. The second comment points out that it is still an open problem whether the correlations predicted by the vacuum state between projections in spacelike related local algebras can be explained by a Reichenbachian common cause located in the intersection of the backward light cones of the regions with which the algebras containing the correlated projections are associated. All we know is that the typical superluminal correlations AQFT predicts possess Reichenbachian common causes located in the union of the backward light cones in question.*

17.1 AQFT

Let me begin by quoting from a recent paper by one of the leading authorities on QFT:

> "So the general framework of AQFT (algebraic quantum field theory) has, for many decades, proved to be consistent with the progress in our theoretical understanding of relativistic quantum physics. It is the appropriate setting for the discussion of the pertinent mathematical structures,

the elaboration of methods for their physical interpretation, the solution of conceptual problems and the classification of theories on the basis of physical criteria." (Buchholz 1998)

I take it that the attitude expressed in this passage is general: AQFT is indeed considered to be an approach to QFT which is sufficiently general to accomodate what is called QFT and which is especially suitable for discussing conceptual issues; so I too take this position here.

The key idea in AQFT is that there is one single entity that characterizes QFT completely: the local net of C^* algebras $V \mapsto \mathcal{A}(V)$ that associates a C^* algebra $\mathcal{A}(V)$ to every open bounded spacetime region V of the Minkowski spacetime M. The interpretation of the local algebra $\mathcal{A}(V)$ is that it (or its selfadjoint part) represents what is *observable* in region V. *All* the information concerning a quantum field theory should be inferred from the net. The properties of the net are prescribed by requiring a few axioms to hold for it. These are: isotony, microcausality, relativistic covariance, existence of vacuum state (together with the spectrum condition, which expresses positivity of energy) and possibly a few more axioms, all physically motivated. More explicitly:

(i) Isotony: if V_1 is contained in V_2, then $\mathcal{A}(V_1)$ is a subalgebra of $\mathcal{A}(V_2)$;
(ii) Einstein (or micro) causality: if V_1 is spacelike separated from V_2, then every element of $\mathcal{A}(V_1)$ commutes with every element of $\mathcal{A}(V_2)$;
(iii) Relativistic covariance: there is a representation α of the identity-connected component $\mathcal{P}$ of the Poincaré group by automorphisms on $\mathcal{A}$ such that $\alpha_g(\mathcal{A}(V)) = \mathcal{A}(gV)$ for all V and $g \in \mathcal{P}$.

The smallest C^* algebra $\mathcal{A}$ containing all the local algebras $\mathcal{A}(V)$) is called the quasilocal algebra. States ϕ are linear functionals defined on this algebra.

(iv) Vacuum representation: for each V, $\mathcal{A}(V)$ is a von Neumann algebra acting on a separable Hilbert space $\mathcal{H}$. There is a distinguished unit vector $\Omega \in \mathcal{H}$, and there is a strongly continuous unitary representation $U(\mathcal{P})$ on $\mathcal{H}$ such that $U(g)\Omega = \Omega$, for all $g \in \mathcal{P}$, and

$$\alpha_g(A) = U(g)AU(g)^{-1}, \quad \text{for all} \quad A \in \mathcal{A}$$

as well as the spectrum condition — the spectrum of the self–adjoint generators of the strongly continuous unitary representation $U(\mathbf{R}^4)$

of the translation subgroup of $\mathcal{P}$ must lie in the closed forward light cone.

The vacuum state ϕ_0 on $\mathcal{A}$ is defined by $\phi_0(A) = \langle \Omega, A\Omega \rangle$, for all $A \in \mathcal{A}$. This state is thus Poincaré invariant: $\phi(\alpha_g(A)) = \phi(A)$, for all $g \in \mathcal{P}$ and $A \in \mathcal{A}$.

(v) Weak additivity: for any non-empty open region V, the set of operators $\bigcup_{g \in \mathbf{R}^4} \mathcal{A}(gV)$ is dense in $\mathcal{A}$ (in the weak operator topology).

See Haag 1992 for further interpretation of these axioms.

17.2 Comment 1: Ontological Silence of AQFT is Not Ontological Neutrality

Two things are remarkable concerning AQFT:

> The hidden richness of these few axioms. One can infer from them deep, non-trivial propositions concerning features of the local net. One sort of results characterizes the structure (the type) of the local algebras, other results concern the algebraic relations implied by the relative positions of the spacetime regions to which they belong (algebraic and statistical independence if the regions are causally independent (spacelike), split inclusion property for regions contained in each other etc.).
> The axioms do not mention *field* or *particle* at all. I call this latter feature of AQFT its *ontological silence.*

In view of the *ontological silence* of AQFT one would think that it is a matter of taste or choice whether one interprets AQFT as a theory about *fields* or about *particles.* But this does not seem to be the case, since AQFT is *not* ontologically *neutral*: While AQFT is compatible with the notion of field, it does *not* seem to be compatible with the notion of particle.

The compatibility is reflected by the fact that we have positive results spelling out the relation of quantum fields (eg. in the sense of the Wightman axioms) to nets of algebras. There are two types of results:

field $\rightarrow$ *net*:

$$\mathcal{A}(V) \equiv \{\Psi[f] : \sup f \subseteq V\}''$$

The idea that a local algebra $\mathcal{A}(V)$ should be generated by the fields smeared by test functions having support in V can be made precise (see eg. Borchers and Yngvason 1992 for precise statements and further references).

The intuitive idea in the *net* $\longrightarrow$ *field* direction is that the field should be obtainable as the "intersection of the local algebras":

net $\longrightarrow$ *field:*

$$\Psi(x) \equiv \bigcap_{V \ni x} \overline{\mathcal{A}(V)}$$

where the "local algebras" $\overline{\mathcal{A}(V)}$ are certain "completions" of the local algebra $\mathcal{A}(V)$, a completion which is needed since the intersection $\bigcap_{V \ni x} \mathcal{A}(V)$ is empty. Results in this direction can be found in Fredenhagen and Hertel 1981.

The incompatibility between AQFT with the notion of particle as an entity which is strictly localized (in spacetime) is a consequence of the *Reeh-Schlieder Theorem*:

Proposition 1 ***(Reeh-Schlieder Theorem)*** *Under the conditions (i)-(v) the vacuum vector Ω is cyclic and separating for $\mathcal{A}(V)$ if V has a non-trivial causal complement.*

It is not immediately clear why the Reeh-Schlieder theorem is a no-go result in connection with a particle interpretation of AQFT; however, the following corollary shows why this is so:

Proposition 2 ***(Corollary of Reeh-Schlieder Theorem)*** *There exists no*

local	$D \in \mathcal{A}(V)$
statistically faithful	$\phi_0(D) = 0$
particle detector observable	$D > 0$

What the corollary shows is that *if* a particle is taken to be an entity that is *strictly* localized in a bounded spacetime region, then this notion of particle is not compatible with the axioms of AQFT. As Haag puts it:

> "... we note that experimentally all information comes from the use of detectors and coincidence arrangements of detectors. The essential features used are that a detector is a macroscopically well localized positive observable which gives no signal in the vacuum state. In the mathematical setup of the theory the two requirements cannot be strictly

> reconciled due to the Reeh-Schlieder theorem; the algebra $\mathcal{A}(V)$ of a strictly finite region does not contain positive operators with vanishing expectation value." (Haag 1992, p. 272]

The standard way out of this conceptual difficulty is to abandon the requirement that a particle is *strictly* localized: approximately localized detector observables and coincidence detectors *can* be modelled in AQFT (see Haag 1992, pp. 272-275).

Another no-go result showing the incompatibility of the notion of strictly localizable particle within the framework of relativistic quantum mechanics is due to D. Malament (1996). Malament shows that under weak conditions there exists no covariant position operator on any spacelike hypersurface Γ in Minkowski spacetime M. To be more precise, consider the system $(\mathcal{H}, a \mapsto U(a), V \mapsto P(V))$, where U is a strongly continuous representation of the translation group in M on the Hilbert space $\mathcal{H}$ and where $\Gamma \ni V \mapsto P(V)$ is the projection measure of the hypothetical position operator of the particle on the hypersurface Γ. For P to be a position operator of a *localizable* particle, one must have

$$\text{localizability:} \quad P(V_1)\ P(V_2) = 0 = P(V_2)\ P(V_1) \quad \text{if } V_1 \cap V_2 = \emptyset \quad (17.1)$$

The relativistic covariance of the position operator P is expressed by the following condition

$$\text{covariance:} \quad P(V + a) = U(a)\ P(V)\ U(-a) \quad (17.2)$$

The following condition expresses the demand that observations in spacelike separated regions are independent:

$$\text{locality:} \quad P(V_1)\ P(V_2) = P(V_2)\ P(V_1) \quad \text{if } V_1 \text{ and } V_2 \text{ are spacelike} \quad (17.3)$$

Proposition 3 ***(Malament's Theorem)*** *If the spectrum of the generator of U is bounded from below, then there exists no P satisfying the localizability, covariance and locality conditions.*

Strictly speaking Malament's Theorem does not apply to AQFT, since in AQFT the region V in the assignment $V \mapsto \mathcal{A}(V)$ is assumed to be an *open* set in M; however, one can show that Malament's Theorem remains valid even if one allows V in $V \mapsto P(V)$ to have a non-zero time width (see Weiner 2001, Proposition III.1.8). Thus, under the hypothesis of Malament's theorem one cannot have a localization operator in AQFT either.

Summary of Comment 1: The ontological significance of the vacuum in AQFT is that its properties indicate that the ontological silence of AQFT is not ontological neutrality: *if* strict localizability (in space or spacetime) is taken as a necessary attribute of a particle, then the properties of the vacuum in AQFT imply that the particle concept is not compatible with AQFT.

17.3 Comment 2: It is Not Known Whether Vacuum Correlations Can Have a Common Cause Explanation

A characteristic feature of AQFT is that it predicts correlations between projections A, B lying in von Neumann algebras $\mathcal{A}(V_1), \mathcal{A}(V_2)$ associated with spacelike separated spacetime regions V_1, V_2 in Minkowski space. Typically, if $\{\mathcal{A}(V)\}$ is a net of local algebras in a vacuum representation, then there exist many normal states ϕ on $\mathcal{A}(V_1 \cup V_2)$ such that $\phi(A \wedge B) > \phi(A)\phi(B)$ for suitable projections $A \in \mathcal{A}(V_1)$, $B \in \mathcal{A}(V_2)$. We call such correlations *superluminal.* The presence of superluminal correlations is one of the consequences of the generic violation of Bell's inequalities in AQFT; specifically, the vacuum state predicts superluminal correlations, since it violates Bell's inequality (see Summers and Werner 1987a, Summers and Werner 1987b, Summers and Werner 1988, Summers 1990 and Halvorson-Clifton 2000 for results concerning the violation of Bell's inequality in AQFT).

According to a classical tradition in the philosophy of science, probabilistic correlations are always signs of causal relations. This is the content of what became called *Reichenbach's Common Cause Principle.* This principle asserts (cf. Salmon 1984) that if two events A and B are correlated, then the correlation between A and B is either due to a direct causal influence connecting A and B, or there is a third event C which is a common cause of the correlation in the sense defined below.

Definition 1 Let A, B be two commuting projections in a von Neumann algebra which are correlated in ϕ:

$$\phi(A \wedge B) > \phi(A)\ \phi(B) \tag{17.4}$$

projection C is a *common cause* of the correlation (17.4) if C commutes with both A and B and the following conditions hold:

$$\frac{\phi(A \wedge B \wedge C)}{\phi(C)} = \frac{\phi(A \wedge C)}{\phi(C)} \frac{\phi(B \wedge C)}{\phi(C)} \tag{17.5}$$

$$\frac{\phi(A \wedge B \wedge C^{\perp})}{\phi(C^{\perp})} = \frac{\phi(A \wedge C^{\perp})}{\phi(C^{\perp})} \frac{\phi(B \wedge C^{\perp})}{\phi(C^{\perp})} \tag{17.6}$$

$$\frac{\phi(A \wedge C)}{\phi(C)} > \frac{\phi(A \wedge C^{\perp})}{\phi(C^{\perp})} \tag{17.7}$$

$$\frac{\phi(B \wedge C)}{\phi(C)} > \frac{\phi(B \wedge C^{\perp})}{\phi(C^{\perp})} \tag{17.8}$$

Definition 1 is a natural specification in a non-commutative probability space of the classical notion of common cause as this was formulated by Reichenbach (1956).

If the correlated projections belong to algebras associated with spacelike separated regions, a direct causal influence between them is excluded by the theory of relativity. Consequently, compliance of AQFT with Reichenbach's Common Cause Principle would mean that for every correlation between projections A and B lying in von Neumann algebras associated with spacelike separated spacetime regions V_1, V_2, there must exist a common cause projection C; however, since observables and hence also the projections in AQFT must be localized in the case of the spacelike correlations predicted by AQFT, one also has to specify the spacetime region V with which the von Neumann algebra $\mathcal{A}(V)$ containing the common cause C is associated. Intuitively, the region V should be disjoint from both V_1 and V_2 but should not be *causally* disjoint from them in order to leave room for a causal effect of C on the correlated events. There are different interpretations of "causal non-disjointness" of V from V_1 and V_2; hence there are different ways to specify a notion of common cause in terms of AQFT. To explore these different concepts, we need some definitions first.

For a point x in the Minkowski space $\mathcal{M}$ let $BLC(x)$ denote the backward light cone of x; furthermore for an arbitrary spacetime region V let $BLC(V) \equiv \bigcup_{x \in V} BLC(x)$. For spacelike separated spacetime regions V_1 and V_2 let us define the following regions

$$wpast(V_1, V_2) \equiv (BLC(V_1) \setminus V_1) \cup (BLC(V_2) \setminus V_2) \tag{17.9}$$

$$cpast(V_1, V_2) \equiv (BLC(V_1) \setminus V_1) \cap (BLC(V_2) \setminus V_2) \tag{17.10}$$

$$spast(V_1, V_2) \equiv \bigcap_{x \in V_1 \cup V_2} BLC(x) \tag{17.11}$$

Region $spast(V_1, V_2)$ consists of spacetime points *each* of which can causally influence *every* point in both V_1 and V_2. Region $cpast(V_1, V_2)$ consists of spacetime points *each* of which can causally influence at least *some* point in *both* V_1 *and* V_2. Region $wpast(V_1, V_2)$ consists of spacetime points *each* of which can causally influence at least *some* point in *either* V_1 *or* V_2.

Obviously it also holds that

$$spast(V_1, V_2) \subseteq cpast(V_1, V_2) \subseteq wpast(V_1, V_2) \tag{17.12}$$

Definition 2 Let $\{\mathcal{A}(V)\}$ be a net of local von Neumann algebras over Minkowski space. Let V_1 and V_2 be two spacelike separated spacetime regions, and let ϕ be a locally normal state on the quasilocal algebra $\mathcal{A}$. If for any pair of projections $A \in \mathcal{A}(V_1)$ and $B \in \mathcal{A}(V_2)$ it holds that if

$$\phi(A \wedge B) > \phi(A)\ \phi(B) \tag{17.13}$$

then there exists a projection C in the von Neumann algebra $\mathcal{A}(V)$ which is a common cause of the correlation (17.13) in the sense of definition 1, then the local system is said to satisfy

Weak Common Cause Principle (WCCP) if $V \subseteq wpast(V_1, V_2)$

Common Cause Principle (CCP) if $V \subseteq cpast(V_1, V_2)$

Strong Common Cause Principle (SCCP) if $V \subseteq spast(V_1, V_2)$

We say that Reichenbach's Common Cause Principle holds for the net (respectively holds in the weak or strong sense) iff for every pair of spacelike separated spacetime regions V_1, V_2 and every normal state ϕ, the Common Cause Principle holds for the local system $(\mathcal{A}(V_1), \mathcal{A}(V_2), \phi)$ (respectively in the weak or strong sense).

If V_1 and V_2 are complementary wedges then $spast(V_1, V_2) = \emptyset$. Since the local von Neumann algebras pertaining to complementary wedges are known to contain correlated projections (see Summers and Werner 1988 and Summers 1990), the *Strong* Reichenbach's Common Cause Principle trivially fails in AQFT.

Problem: Does Reichenbach's Common Cause Principle hold in quantum field theory ?

The above question was first formulated in Rédei 1997 (see also Rédei 1998) and the answer to it is not known. What is known is that the *Weak* Reichenbach's Common Cause Principle typically holds under mild assumptions on the local net $\{\mathcal{A}(V)\}$:

Proposition 4 *If a net $\{\mathcal{A}(V)\}$ with the standard conditions (isotony, Einstein locality, Poincaré covariance, weak additivity, spectrum condition) is such that it also satisfies the local primitive causality condition and the algebras pertaining to double cones are type III, then every local system $(\mathcal{A}(V_1), \mathcal{A}(V_2), \phi)$ with V_1, V_2 contained in a pair of spacelike separated double cones and with a locally normal and locally faithful state ϕ satisfies Weak Reichenbach's Common Cause Principle.*

(See Rédei and Summers 2002 for the proof of the above proposition and for additional analysis of the status of Reichenbach's Common Cause Principle in quantum field theory.)

Local primitive causality is a condition that expresses the hyperbolic character of time evolution in AQFT. For a spacetime region V let $V'' = (V')'$ denote the *causal completion* (also called *causal closure* and *causal hull*) of V, where V' is the set of points that are spacelike from every point in V. The net $\{\mathcal{A}(V)\}$ is said to satisfy the *local primitive causality* condition if $\mathcal{A}(V'') = \mathcal{A}(V)$ for every nonempty convex region V. Local primitive causality is a condition that is known *not* to hold for some nets of local algebras satisfying the standard axioms (Garber 1975); however, this condition has been verified in many concrete models.

Summary of Comment 2: The vacuum state predicts superluminal correlations, i.e. correlations between projections belonging to von Neumann algebras associated with spacelike separated spacetime regions V_1 and V_2. Such correlations should be explainable by a common cause lying in the intersection of the backward light cones of V_2 and V_2. It is not known whether such common causes exist or not; however, it is known that common causes of superluminal correlations predicted by faithful states lying in the *union* of the backward light cones of V_1 and V_2 do exist (Proposition 4). This proposition indicates that AQFT is a causally rich theory; in particular it indicates that AQFT is a theory that possibly complies with Reichenbach's Common Cause Principle; yet this is yet to be proved.

Acknowledgements

Work supported by OTKA (contract numbers: T 032771, T 024841 and T 025880).

References

H-J. Borchers and J. Yngvason (1992): From quantum fields to local von Neumann algebras. *Reviews in Mathematical Physics*, Special Issue 15.

D. Buchholz (1998): Current trends in axiomatic quantum field theory. Talk at "Ringberg Symposium on Quantum Field Theory", Ringberg Castle.

K. Fredenhagen and J. Hertel (1981): Local algebras of observables and pointlike localized fields. *Communications in Mathematical Physics*, 80: 555–561.

W.D. Garber(1975): The connexion of duality and causality properties for generalized free fields. *Communications in Mathematical Physics* 42: 195–208

R. Haag (1992): *Local Quantum Physics. Fields, Particles, Algebras.* Berlin: Springer Verlag.

H. Halvorson and R. Clifton (2000): Generic Bell correlations between arbitrary local algebras in quantum field theory. *Journal of Mathematical Physics*, 41: 1711–1717.

D. Malament (1996): In defense of dogma: Why there cannot be a relativistic quantum mechanics of (localizable) particles. In R. Clifton (ed.), *Perspectives on Quantum Reality. Non-relativistic, Relativistic and Field Theoretic*, Dordrecht, Boston and London: Kluwer Academic Publishers, 1–10.

M. Rédei (1997): Reichenbach's common cause principle and quantum field theory. *Foundations of Physics*, 27: 1309–1321.

M. Rédei (1998): *Quantum Logic in Algebraic Approach.* Dordrecht, Boston and London: Kluwer Academic Publishers.

M. Rédei and S.J. Summers (2002): Local primitive causality and the common cause principle in quantum field theory. *Foundations of Physics*, 32: 335-355

H. Reichenbach (1956): *The Direction of Time.* Los Angeles: University of California Press.

W.C. Salmon (1984): *Scientific Explanation and the Causal Structure of the World.* Princeton: Princeton University Press.

S.J. Summers and R. Werner (1985): The vacuum violates Bell's inequalities. *Physics Letters A*, 110: 257–279.

S.J. Summers and R. Werner (1987a): Maximal violation of Bell's inequalities is generic in quantum field theory. *Communications in Mathematical Physics*, 110: 247–259.

S.J. Summers and R. Werner (1987b): Bell's inequalities and quantum field theory I. General setting. *Journal of Mathematical Physics*, 28: 2440–2447.

S.J. Summers and R. Werner (1988): Maximal violation of Bell's inequalities for algebras of observables in tangent spacetime regions. *Annales de l'Institut Henri Poincaré – Physique théorique*, 49: 215–243.

S.J. Summers (1990): On the independence of local algebras in quantum field theory. *Reviews in Mathematical Physics*, 2: 201–247.

M. Weiner (2001): Lokalizáció és Részecskekép ("Localization and the Particle Picture"). Masters Thesis, ELTE, Budapest.

This page is intentionally left blank

List of Contributors

Sunny Y. Auyang
100 Memorial Drive #521B, Cambridge, MA 02142, USA
E-mail: sauyang@prodigy.net

Jeffrey A. Barrett
Department of Logic & Philosophy of Science, School of Social Science,
University of California at Irvine, Irvine, CA 92697, USA
E-mail: jabarret@uci.edu

Rob Clifton
Department of Philosophy, 1001 Cathedral of Learning,
University of Pittsburgh, Pittsburgh, PA 15260, USA
E-mail: rclifton@pitt.edu

Dennis Dieks
History and Foundations of Science, State University of Utrecht,
P.O. Box 8000, NL-3508 TA Utrecht, The Netherlands
E-mail: D.G.B.J.Dieks@phys.uu.nl

Michael Drieschner
Institut für Philosophie, Ruhr-Universität-Bochum,
D-44780 Bochum, Germany
E-mail: Michael.Drieschner@ruhr-uni-bochum.de

Tim Oliver Eynck
NIKHEF, Postbus 41882, NL-1009 DB Amsterdam, The Netherlands
E-mail: teynck@nikhef.nl

Brigitte Falkenburg
Universität Dortmund, FB 14: Fach Philosophie,
Emil-Figge-Str. 50, D-44227 Dortmund, Germany
E-mail: falkenburg@fb14.uni-dortmund.de

Gordon N. Fleming
Pennsylvania State University, 325 Davey Lab,
University Park, PA 16802, USA
E-mail: gnf1@earthlink.net

Hans Halvorson
Department of Philosophy, Princeton University,
Princeton, NJ 08544, USA
E-mail: hhalvors@princeton.edu

Nick Huggett
Department of Philosophy – MC 267, University of Illinois at Chicago,
601 South Morgan Street, Chicago, IL 60607-7114, USA
E-mail: huggett@uic.edu

Meinard Kuhlmann
Universität Bremen, FB 9: Studiengang Philosophie,
Postfach 33 04 40, D-28334 Bremen, Germany
E-mail: mkuhlmann@uni-bremen.de

Holger Lyre
Philosophisches Seminar (LFB 1), Universität Bonn,
Am Hof 1, D-53113 Bonn, Germany
E-mail: lyre@uni-bonn.de

Miklós Rédei
Department of History and Philosophy of Science, Faculty of Natural Sciences, Eötvös University, P.O. Box 32, H-1518 Budapest, Hungary
E-mails: redei@ludens.elte.hu, redei@hps.elte.hu

Michael Redhead
Centre for Philosophy of Natural and Social Science,
London School of Economics, and Political Science,
Houghton Street, UK-London WC2 2AE, United Kingdom
E-mail: mlr1000@cus.cam.ac.uk

Simon Saunders
University of Oxford, Sub-Faculty of Philosophy,
10 Merton Street, UK-Oxford OX1 4JJ, United Kingdom
E-mail: simon.saunders@philosophy.oxford.ac.uk

Johanna Seibt
University of Aarhus, Department of Linguistics,
DK-8000 Aarhus C, Denmark
E-mail: seibt@mail.utexas.edu

Peter Simons
School of Philosophy, University of Leeds,
UK-Leeds LS2 9JT, United Kingdom
E-mail: p.m.simons@leeds.ac.uk

Paul Teller
Philosophy Department, University of California at Davis,
Davis, CA 95616, USA
E-mail: prteller@ucdavis.edu

Andrew Wayne
Department of Philosophy, Concordia University,
1455 de Maisonneuve Blvd. West, Montreal, Quebec H3G 1M8, Canada
E-mail: awayne@alcor.concordia.ca

Index

Ontological Aspects of Quantum Field Theory

Quantum field theory (QFT) provides the framework for many fundamental theories in modern physics, and over the last few years there has been growing interest in its historical and philosophical foundations. This anthology on the foundations of QFT brings together 15 essays by well-known researchers in physics, the philosophy of physics, and analytic philosophy.

Many of these essays were first presented as papers at the conference "Ontological Aspects of Quantum Field Theory", held at the Zentrum für interdisziplinäre Forschung (ZiF), Bielefeld, Germany. The essays contain cutting-edge work on ontological aspects of QFT, including: the role of measurement and experimental evidence, corpuscular versus field-theoretic interpretations of QFT, the interpretation of gauge symmetry, and localization.

This book is ideally suited to anyone with an interest in the foundations of quantum physics, including physicists, philosophers and historians of physics, as well as general readers interested in philosophy or science.

ISBN 981-238-182-1

9 789812 381828

www.worldscientific.com

5117 hc