Ships, Lifts and Cages

A unified interpretation of transformations and symmetry in classical and quantum physics

MSc Thesis

written by

Nicholas P.C. Evans

under the supervision of **Prof. Dr. D.G.B.J. Dieks**

at the *Universiteit Utrecht*

August 2012

Universiteit Utrecht

Institute for History and Foundations of Science (IGG)

Acknowledgments

I would like to thank my supervisor Dennis Dieks for his support throughout this project and for his careful comments on my work. I would also like to thank him for supporting my application to present a preliminary version of this work at the COST conference *Fundamental Problems in Quantum Mechanics* in Malta. That was an inspirational experience for me.

On a personal level, I would like to thank Peter, Jenny and Benj for their support, as well as Sara, An and Mart.

And thank you in advance to the reader, I did my best to make things clear. Needless to say, I will not always have succeeded.

Abstract

In this thesis, I explore the extent to which a focus on transformations and symmetry can highlight the conceptual unity of classical spacetime theories and quantum (gauge) theory, in particular general relativity (GR) and non-relativistic quantum-electromagnetism (Q-EM). In the literature, there is a general belief that because the transformations of spacetime theories act on external space, they have a different interpretation from the transformations of gauge theory, which act on internal space. By focusing both on the role of transformations within these different theories, and the possibility of interpreting the transformations actively, I show that, despite important differences, there is more conceptual unity than is often realised. Specifically, I aim to answer the two following questions.

1) *What role do symmetry transformations play in GR and Q-EM?* I show that the equivalence principle and the gauge principle play analogous roles in GR and Q-EM respectively. This leads to an analogue of the equivalence principle in Q-EM, and the formulation of the notion of an inertial gauge. The claim that "dynamical forces restore symmetry" is demystified.

2) *Is there a fundamental difference between spacetime transformations and gauge transformations?* It is commonly accepted that gauge transformations, unlike spacetime transformations, cannot be interpreted actively. I show why this is mistaken by considering the action of gauge transformations on both classical and quantum systems. In the latter case, I suggest that the particular nature of quantum systems is responsible for the differences that there appear to be between active gauge transformations and active spacetime transformations.

Contents

Chapter 1

Situations of symmetry

It is the July of 1952 and, at the conference "Symposium on New Research Techniques in Physics" in Rio de Janeiro, Eugene Wigner is engaged in a discussion on the nature of symmetries and conservation laws in physics.[1] A participant, Mr Medina, asks Wigner whether one of the previous questions he has just answered was related to the "gauge invariance of charge fields under rotation of the complex field function". Wigner answers that it was. At this point, Bohm intervenes

> I would just like to ask prof. Wigner if he has any speculative ideas as to whether there is some geometric or mechanical transformation which corresponds to this symmetry?

Wigner answers as follows

> I think none beyond the point which dr. Medina just so well described [that the total charge operator is proportional to the infinitesimal rotations of charge space]. I think that we should admit that we do not have an understanding of the deeper causes of any dynamic symmetry. There seems to be an analogy suggested by experimental fact [the nature of this analogy isn't clear]. However, an explanation in the same sense as we have an explanation, for instance, for the hydrogen spectrum, is entirely absent. We see only connections but not more than that.

In Bohm's question and Wigner's answer are the main ingredients of this thesis. Bohm's question can be interpreted as an inquiry into the possibility of active gauge transformations. Wigner's negative response is representative of the view, common to this day, that gauge transformations cannot be

[1]This anecdote is based on the transcript of a conference discussion between Eugene Wigner, Tiomno, Leite Lopes, Medina and Bohm (who I assume must be David Bohm) [Wigner, 1992b, p. 107].

interpreted actively [Brading and Brown, 2004]. In chapter 5 I will question the validity of this conclusion. Wigner's further comment on the "causes of dynamic symmetry" refers to the mystery surrounding the role of the so-called gauge principle in apparently allowing the existence of dynamical forces (such as electromagnetism) to be derived from mere symmetry requirements. The methodological role of symmetry transformations will be discussed in chapter 4. This exchange is also interesting for historical reasons. Seven years later, in 1959, Bohm would publish, together with Aharonov, the seminal paper *Significance of Electromagnetic Potentials in Quantum Theory*, in which they predict the Aharonov-Bohm (AB) effect [Aharonov and Bohm, 1959]. This prediction, later well confirmed by experiment, has significant consequences both for the possibility of active gauge transformations, and for understanding the role of dynamic symmetries (as they are called by Wigner). It is thus interesting to note that such issues had been on Bohm's mind for quite some time before the publication of the famous paper with Aharonov. A review of the AB effect will be given in chapter 4.

The first three chapters will do the groundwork necessary for fully understanding the more technically involved final two chapters. Chapter 1 introduces the notion of a transformation in general, and presents some of the difficulties that arise in attempting to formulate a satisfactory notion of symmetry in this context. To some extent, these difficulties will be resolved in chapter 5. Chapter 2 introduces the notion of a gauge transformation, and places it in the context of Wigner's canonical philosophy of symmetry. The influence of Wigner's views on more recent work in the philosophy of physics will also be discussed. Chapter 3 is devoted to the exposition and definition of important mathematical notions. Some of these, for instance the notion of a coordinate system, will appear in earlier chapters. However, it will be assumed that, until chapter 3, an intuitive grasp of such concepts is sufficient.

By the end of the thesis, I hope to have answered to the two following questions:

1. *What role do symmetry transformations play in general relativity and non-relativistic quantum-electromagnetism?*

2. *Is there a fundamental difference between spacetime transformations and gauge transformations?*

Ultimately, underlying these two questions is a third one about the unity of physics. The unification of the four forces (gravitational, electromagnetic, strong and weak) is one of the main motivations driving today's theoretical physics. By showing that, with respect to transformations and symmetry, a certain amount of unity already exists, I hope to make a very modest contribution in this direction. Furthermore, by suggesting where the unity breaks down, I hope to point towards possible areas of future research.

1.1 The ship, the cage and the lift

The word "symmetry" is omnipresent in modern physics. However, one should not confuse the word itself with its use in a particular context. As Brading and Castellani aptly remark, the notion of "symmetry" is so ubiquitous in modern physics that it is impossible to provide a unified general account of its role in today's theories [Brading and Catellani, 2003, p.11]. Pierre Curie used a notion of symmetry to study the properties of crystals. Einstein relied on a notion of symmetry to arrive at his theories of special and general relativity. Emmy Noether derived a relation between symmetry and conservation laws. At the LHC, they are exploring the implications of spontaneous symmetry breaking. Van Fraassen points out that symmetry principles can play a crucial role in problem solving [van Fraassen, 1989]. Weyl and Wigner were pioneers of the use of symmetry (expressed in the mathematical language of group theory) in quantum physics. Any research into the role of symmetry in modern physics must therefore start by determining precisely in which context the subject is to be studied.

The general notion of symmetry that we will be concerned with is that of "invariance under a transformation". Two questions immediately arise: Invariance of *what*? (2) Transformed *how*? In other words, given that we start with a "something" with certain features, we need to know how this something is changed (which transformation is performed) and what it is about this something that does not change (which features are invariant). The "something" of relevance here is either a given physical situation, or a mathematical representation thereof. In the first case, the notion of invariance refers to the fact that the transformed physical situation is, in a way to be made precise, indistinguishable from the untransformed situation. However, if the transformation does actually do something to the physical situation, then the transformed and untransformed situations must also differ in some way. As we will see, balancing these two contradictory requirements is a delicate but crucial aspect of any interpretation of symmetry. When a transformation acting on a physical situation satisfies these requirements, I will call it an "empirical symmetry" [Healey, 2009].

In the second case, when the transformation acts on a mathematical representation of a physical situation, the notion of invariance refers to the fact that the transformed and untransformed situations both satisfy the equations of the theory. I will call these "theoretical symmetries" [Healey, 2009]. I now turn to a more detailed analysis of these two different kinds of symmetry, using Healey's definitions as starting points.

1.1.1 Empirical symmetries

Healey gives the following definition [Healey, 2009, p. 703]

Empirical symmetry (Healey): A bijective map $\phi : S \to S$ of a set of situations onto itself is an *empirical symmetry with respect to C-type measurements* if and only if no two situations related by ϕ can be distinguished by measurements of type C.

A first thing to note about this definition is that it makes the identity an empirical symmetry. This is a trivial result we would like to exclude. This can be achieved by adding some additional requirement such as "for all $s \in S$, $\phi(s) \neq s$". In other words, we require that ϕ represent a transformation which actually changes something about s. We will see later that this "transformation condition" becomes a major point of discussion in the literature.

A second problematic feature of this definition is the vague nature of the "set of situations", S. Elements $s \in S$ are "actual physical situations", and this makes them hard to handle [Healey, 2009, p. 702]. Healey specifies that S should include "actual" as well as "possible" situations [Healey, 2009, p. 703]. However, the world is not something that lends itself to being easily carved up into well defined situations. As an example of the kind of elements one might expect to find in S, Healey mentions the case of "purely mechanical phenomena in a Newtonian world" [Healey, 2009, p. 704]. Unfortunately, the precision of this example is achieved at the expense of blurring the distinction between empirical and theoretical symmetries. Actual physical situations do not take place in Newtonian worlds, or in any kind of theoretical world for that matter. They just occur out-there. In fact, I do not believe that the elements of S can be determined in an unambiguous way. As we will see later, this is not the case for theoretical symmetries, where the elements of the set of (mathematical) models of a theory can be sharply and unambiguously defined.

To address this concern, I suggest abandoning the notion of a "set of situations". Instead, I will define "empirical symmetries" on a case by case basis.

Empirical symmetry: A transformation T that acts on a physical situation s is an empirical symmetry with respect to C-type measurements, if and only if $T(s)$ and s cannot be distinguished by measurements of type C and T obeys the transformation condition $T(s) \neq s$.

By removing the dependency on a "set of situations", this definition implicitly acknowledges that the exact nature of the "physical situation" in question will depend the particular application of the definition. Furthermore, not all cases to which this definition can be applied will be equally interesting. For instance, imagine that s is the current state of the solar system and $T(s)$ is the same as s in all respects except that a star at the other end of the universe has been displaced by a few nanometers. Clearly, s and $T(s)$ will be indistinguishable by all measurement apparatuses currently available. However, this empirical symmetry is not surprising and does not have any significant ramifications for our scientific understanding of the world. In a similar vein, one could imagine two situations s and $T(s)$ which are quite clearly different, but that could be made indistinguishable by imposing extreme restrictions on the type of measurements allowed. The upshot of these considerations is the realization that a certain amount of discretion is necessary to identify those transformations that are actually interesting. As the following examples will show, empirical symmetries worth thinking about surprise our everyday physical intuitions.

Particularly interesting empirical symmetries are ones for which the transformed situations cannot be distinguished by experiments "confined to the situation" (exactly what this means will be clarified by the upcoming examples) [Healey, 2009, p. 703]. If an empirical symmetry holds with respect to experiments of such type, we call it, following Healey, a "*strong empirical symmetry*". I will now give three examples of empirical symmetries. The first two, Galileo's ship and Faraday's cage, are strong empirical symmetries. The third, Einstein's lift, differs from the first two because it is only an approximate empirical symmetry, and therefore can be detected by very careful experiments confined to the situation. The full significance of Einstein's lift will become evident in chapter 4 when we discuss the role of the equivalence principle in general relativity.

Galileo's ship

In an often-quoted passage, Galileo describes the following situation [Galileo, 1967, p. 187]

> Shut yourself up with some friend in the main cabin below decks on some large ship, and have with you there some flies, butterflies, and other small flying animals. Have a large bowl of water with some fish in it; hang up a bottle that empties drop by drop into a wide vessel beneath it. With the ship standing still, observe carefully how the little animals fly with equal speed to all sides of the cabin. The fish swim indifferently in all directions; the drops fall into the vessel beneath; and, in throwing

> something to your friend, you need throw it no more strongly in one direction than another, the distances being equal; jumping with your feet together, you pass equal spaces in every direction. When you have observed all these things carefully (though there is no doubt that when the ship is standing still everything must happen in this way), have the ship proceed with any speed you like, so long as the motion is uniform and not fluctuating this way and that. You will discover not the least change in all the effects named, nor could you tell from any of them whether the ship was moving or standing still. In jumping, you will pass on the floor the same spaces as before, nor will you make larger jumps towards the stern than toward the prow even though the ship is moving quite rapidly, despite the fact that during the time that you are in the air the floor under you will be going in a direction opposite to your jump. In throwing something to your companion, you will need no more force to get it to him whether he is in the direction of the bow or the stern, with yourself seated opposite. The droplets will fall as before into the vessel beneath without dropping towards the stern, although while the drops are in the air the ship runs many spans. The fish in their water still swim toward the front of their bowl with no more effort than toward the back, and will go with equal ease to bait placed anywhere around the edges of the bowl. Finally, the butterflies and flies will continue their flight indifferently toward every side, nor will it happen that they are concentrated toward the stern, as if tired out from keeping up with the course of the ship, from which they will have been separated during long intervals by keeping themselves in the air. And if smoke is made by burning some incense, it will be seen going up in the form of a little cloud, remaining still and moving no more toward one side than the other. The cause of all these correspondences of effects is the fact that the ship's motion is common to all the things contained in it, and to the air also.

Although strictly speaking Galileo's ship is not in inertial motion because it is in circular motion around the center of the earth, the example is generally interpreted as illustrating the principle of Galilean relativity. This states the following:

Principle of Galilean relativity: Two copies of the same system in inertial motion with respect to one another will evolve in empirically indistinguishable ways.

As Galileo does in the passage quoted above, one of the best ways to understand how two (different) physical situations can be related by an empirical symmetry is to phrase the transformation in terms of an observer constrained to make measurements on a limited environment. In the case of Galileo's ship, we are invited to consider the observations that someone "trapped" in the cabin below deck could make on the objects in his immediate surroundings. The transformation is a symmetry if the observer, by appealing to his limited experiments, cannot tell if it has taken place.[2] We might sum this up in the slogan form "the observer cannot tell the difference".

The limits imposed on the measurement capacities of the observer clarify what Healey referred to as measurements "confined to the situation". Indeed, since the transformation does actually change the state of motion of the ship, if no constraints were imposed, the observer could quite easily tell if the ship was at rest in the harbour or sailing at a uniform velocity away from it. All he would have to do is open the window and look at the shore, or, if he is more sophisticated, determine the speed of the boat with respect to the shore by means of RADAR pulses.

In terms of our formalism, Galileo's ship can be summarized in the following terms. The situations s and $T(s)$ are the mechanical states of various objects and animals in a cabin below deck of a large ship. The transformation T changes the ship from one state of uniform motion (with respect to some reference point) to a different state of uniform motion (with respect to the same reference point).

Faraday's cage

Michael Faraday gives another example of a transformation that is a strong empirical symmetry. He describes building a hollow cube (also referred to as Faraday's cage) large enough for him to fit inside, covering it with a good conducting material and insulating it well from the ground. He then proceeds to charge the conducting material to such an extent that "sparks flew from its surface". He continues as follows [Healey, 2009, p. 699]

> I went into this cube and lived in it, but though I used lighted candles, electrometers, and all other tests of electrical states, I could not find the least influence on them.

Charging the surface of a hollow conductor results in a uniform increase of the electric potential $A_0 = \phi$ inside the conductor, while the magnetic vector

[2] This appeal to what an observer can measure assumes that the transformation T is "adiabatic". This means that it happens slowly and gently enough not to disturb the elements of the environment in which the observer is situated. If T was not adiabatic, then disturbances due to the violence of the transformation would break the symmetry, even though these disturbances are not, strictly speaking, consequences of performing the transformation T in itself.

potential $\mathbf{A}$ stays unchanged. If the electromagnetic 4-vector potential A_μ vanished inside the conductor before the transformation, then the $\mathbf{E}$ and $\mathbf{B}$ fields given by

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{1.1}$$

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} \tag{1.2}$$

will also vanish after the transformation. Thus, charging the surface of the conductor results only in bringing the inside of the conductor to a higher potential, without inducing any change in electric and magnetic fields. Faraday could find no difference in the outcomes of the experiments he performed within the charged hollow conductor, and therefore concluded what one could call the principle of Faradean relativity. This could be stated (in a similar form to the Galilean version) as follows:

Principle of Faradean relativity: Two copies of the same system at different uniform electric potentials will evolve in empirically indistinguishable ways.

As we will see in chapter 4, a uniform change of electric potential in a region of space does have an effect on the phase of the wavefunction of a quantum particle in that region. The Aharonov-Bohm effect shows that such a phase shift can have empirical consequences. However, whether or not this poses problems for the principle of Faradean relativity is a subtle issue that turns on the precise definition of a "system" that one adopts. These issues will be dealt with in chapter 5.

Within the domain of classical physics, the principle can be accepted unproblematically. Once again, the restrictions on the measurements that the observer inside Faraday's cage is allowed to make play an important role. If the observer is allowed to somehow compare his situation with that of an observer outside the cage at a lower potential (for instance by measuring the potential difference between the two with a voltmeter), then he will be able to tell that he is at a higher potential. 't Hooft describes this possibility vividly in terms of the fate of a squirrel who has one foot on an electrified power line and the other on a grounded conductor ['t Hooft, 1980, p. 97]. However, as long as the squirrel keeps both feet on the power line, it will suffer no harm.

In terms of our formalism, Faraday's cage can be summarized as follows. The situations s and $T(s)$ are the mechanical and electromagnetic states of some (non-quantum) objects inside a hollow conducting container. The transformation T charges the surface of the container, thereby raising the electric potential inside the container uniformly (no spatial dependence) from its previous value to some new value.

Einstein's lift

The third empirical symmetry is that of Einstein's lift, or in a different form, Einstein's freely falling man, both famous as illustrations of the equivalence principle [Norton, 1985, p. 204]. In the latter case, the claim is that a man cannot distinguish being in a state of inertial motion, or being in a state of non-inertial motion in a gravitational field. Einstein writes [Janssen, 2011, p. 4]

> Because for an observer in free-fall from the roof of a house, there is during the fall, at least in his immediate vicinity, no gravitational field. Namely, if the observer lets go of any bodies, they remain, relative to him, in a state of rest or uniform motion, independent of their special chemical or physical nature. The observer, therefore, is justified in interpreting his state as being "at rest".

It may seem confusing that in this statement, Einstein writes that, in the vicinity of the falling man, there is "no gravitational field". Of course, it is precisely because there *is* a gravitational field that the man, who is actually in a state of non-uniform motion with respect to the house, feels as if there is no gravitational field... The source of this confusion over what exactly counts as a "gravitational" field will be addressed in detail in chapter 4. The lift version of the symmetry goes as follows. Consider a person in a lift with no windows. This person feels a force against her feet. She cannot know if the lift is sitting on the surface of the earth, and the force is due to the gravitational field of the earth, or if she is in outer-space, far from any gravitational influence, and there are rocket boosters attached to the lift causing her to accelerate.

In these cases, the role of limited measurements once again plays an important role. For instance, if the observer is allowed to look out of the lift, she can see if there are any massive bodies nearby that might be responsible for the force she feels, and she can also check if there are rocket boosters attached to the lift. However, in contrast to Galileo's ship and Faraday's cage, in Einstein's lift it is also possible to tell the transformation has taken place with experiments confined to the situation, provided these are accurate enough. If the observer has a large enough lift and good enough instruments, she might try to detect the presence of tidal forces acting on objects inside the lift. Thus, if the lift is on the surface of the earth, then two balls dropped from the top of the lift will be very slightly closer together when they reach the bottom. However, if she is in outer space, the distance between the two balls will stay constant, from the time they are released near the top of the lift to the time they hit the floor. Thus, the symmetry holds as long as the observer is not allowed to look for "sources" of the forces she feels, and as long as she limits her measurements to small enough distances and short enough times.

I will not try and formulate a principle corresponding to this empirical symmetry because that would amount to formulating the equivalence principle, and the various difficulties involved in achieving this will be covered in chapter 4. The difficulty in formulating such a principle further underlines the difference between this empirical symmetry and the two others. Nevertheless, the principle can still be summarized using our formalism. s and $T(s)$ are the states of motion of various different bodies in the elevator. T involves one of either: (1) introducing a source of gravitation (some stress-mass-energy density) and allowing the lift to free-fall (2) removing a source of gravitation and turning on rocket boosters attached to the lift.

In all three cases of empirical symmetry just discussed, there is no problem in determining that the transformation condition, $T(s) \neq s$, is satisfied. As was shown, the observer "in the situation" can always tell if the transformation has taken place when there are no restrictions placed on the experiments he is allowed to perform. Furthermore, for an observer "outside the situation" it is always obvious whether the transformation has taken place. This is because the transformation only acts on a subsytem of the universe, and the physical state of the outside observer is not affected by it. It is the change in relations between the outside observer and the system on which the transformation acts that enables the transformation condition to always be satisfied in these cases [Brading and Brown, 2004].

1.1.2 Theoretical symmetries

In the previous section we presented several examples of how physical situations could be transformed without an observer, contained within the situations, being able to tell that the transformation had taken place. In this section, we will consider how transformations act on the mathematical structures that physicists use to represent the physical world. Since physics deals mainly in such mathematical structures, the analysis of "theoretical symmetries" will be our main concern throughout the rest of this thesis. Nevertheless, the physical interpretation of the transformations between such mathematical structures will be of the utmost importance. We will want to know whether the new mathematical structure obtained after a particular transformation corresponds to a different description of the same physical situation, or whether it corresponds to a different physical situation altogether. This concern did not arise as such in the discussion of empirical symmetries. In those cases, it was always clear what the result of the transformation was. When we move away from the intuitive realm of physical situations into the more abstract realm of their mathematical representations, we must be aware that not all changes in representation correspond to changes in the world. One of the main aims of this thesis is to arrive at an understanding of the significance of transformations that take place in the mathematical domain. We will say that a transformation between mathematical structures has *direct empirical significance* when the structures that it relates can be interpreted as representing physically different situations. We will call these *active* transformations. However, we will also argue that transformations that correspond simply to a re-description of the same physical situation also have a central role in physical theories. We will call these *passive* transformations.

In the rest of this chapter I will introduce the kind of issues with which we will be confronted in the rest of this thesis by taking examples from the simple context of Newtonian mechanics in its original (non-generally covariant) formulation. Let us start with Healey's definition of a theoretical symmetry [Healey, 2009, p. 706]

Theoretical symmetry (Healey): A mapping $f : M \to M$ of the set of models of a theory Φ on to itself is a *theoretical symmetry* if and only if the following condition holds: For every model m of Φ that may be used to represent (a situation s in) a possible world w, $f(m)$ may also be used to represent (s in) w. Two models related by a theoretical symmetry of Φ are *theoretically equivalent in* Φ.

A model of a theory is to be understood as a mathematical situation (to be distinguished from the physical situations of the previous section) in which the mathematical objects to which the theory applies obey the laws (equations of motion) of the theory. Some philosophers like to define the concept of a *theory* itself in terms of the set of models that it allows (see for instance [Ismael and van Frassen, 2003, p. 372]). Whether or not this is a good move is not relevant to this discussion. The important issue is that a given theory restricts the set of authorized mathematical situations. Since these situations are mathematically described, they are sharply defined, and it is not problematic to talk of the set of models of a theory.

As in the case of empirical symmetries, not all theoretical symmetries are equally interesting. Ismael and van Frassen devote considerable attention to finding criteria to identify the interesting transformations between models of a theory [Ismael and van Frassen, 2003]). However, my focus in this thesis will be on the significance of specific transformations, and therefore I will not address the problem of finding a general way to distinguish interesting theoretical symmetries from uninteresting ones.

I will now focus on one example of a transformation that takes place in the mathematical domain, and use it to highlight the kinds of problem we will be faced with in the upcoming chapters. Take the trajectory of a single massive particle in an empty universe that obeys Newton's laws of motion. This is a model of Newtonian mechanics. Let us represent the trajectory of the particle by a 3-vector $\mathbf{x}(t)$ in a real vector space $\mathbb{R}^3$ that is a function of (Newtonian absolute) time t. If there are no forces acting on the particle, then the trajectory of the particle is a model of Newtonian mechanics if, for all t and for each component x_i of $\mathbf{x}$

$$\frac{\mathrm{d}^2 x_i}{\mathrm{d}t^2} = 0 \tag{1.3}$$

I can generate another model of the theory by performing the following transformation

$$\mathbf{x}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} \to f(\mathbf{x}(t)) = \mathbf{x}'(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix} = \begin{pmatrix} x(t) + v_x t \\ y(t) \\ z(t) \end{pmatrix} \tag{1.4}$$

where v_x is a constant in time. The transformed trajectory $f(\mathbf{x}(t))$ satisfies (1.3) for all t, and is therefore also a model of Newtonian mechanics. At this point, most authors would already accept that f is a symmetry transformation in Newtonian mechanics. For instance, Brading and Brown write [Brading and Brown, 2004, p. 645]

> A symmetry transformation is a transformation of these variables [the dependent and independent variables] that preserves the explicit form of the laws.

In this case, the "form of the laws" is given by (1.3) and the transformation of the variables by (1.4). When the second is plugged into the first, the extra term $v_x t$ is killed by the second order time derivative, and the equation returns to its original form.

However, according to Healey's definition, f is only a theoretical symmetry if $f(\mathbf{x}(t))$ and $\mathbf{x}(t)$ may both "be used to represent (s in) w". This seems to imply that both $f(\mathbf{x}(t))$ and $\mathbf{x}(t)$ must represent the *same* physical situation s. In order to determine whether this is the case, it is necessary to have a physical interpretation of the transformation f, over and above its mathematical definition.

In fact, there are two ways in which f can be interpreted. The first is as a change in description of the same physical situation called a coordinate transformation. After the transformation, we measure distance in the x

direction with respect to a point which, in the first coordinate system, moves in the negative x direction at speed v_x. Thus, a point with coordinates $x(t)$ in the first coordinate system will have coordinates $x'(t) = x(t) + v_x t$ in the second (if, at $t = 0$, $x'(t) = x(t)$). Because this interpretation is possible, $f(\mathbf{x}(t))$ and $\mathbf{x}(t)$ do both represent the same physical situation s, and therefore f is a symmetry according to Healey. Since this transformation is interpreted as a change in description of the same physical situation, it is a *passive* transformation.

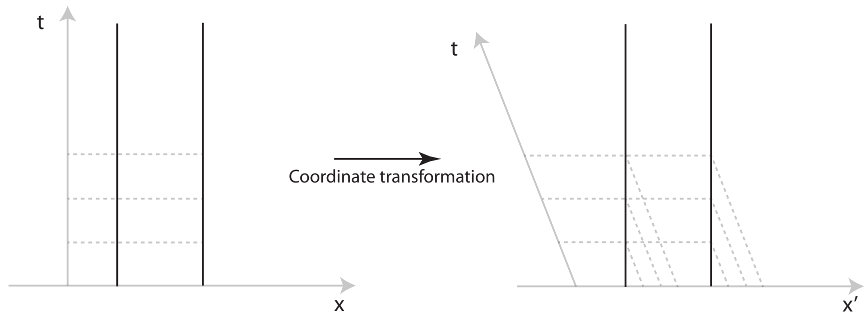

Figure 1.1: Representation of a coordinate transformation in Galilean spacetime, where the worldlines could represent the front and back of Galileo's ship. After the transformation the ship appears to be in motion according to the x' coordinate.

However, there is another way to interpret the above transformation. $f(\mathbf{x}(t))$ might give the trajectory of some other particle, which moves with speed v_x in the x direction with respect to the particle described by $\mathbf{x}(t)$ (similarly, it might represent setting the original particle into a state of uniform motion). In this case, $f(\mathbf{x}(t))$ and $\mathbf{x}(t)$ do *not* represent the same physical situation. However, they do both represent models of Newtonian mechanics, because the motion of both particles obeys the equation (1.3). Since this transformation is interpreted as generating a different physical situation, it is an *active* transformation.

The above considerations suggest that any transformation f of the dependent and independent variables has a dual interpretation, either as a passive or as an active transformation. Conditions for the transformation to count as a symmetry can be given for both. As in the case of empirical symmetries, I prefer to formulate these conditions in terms of individual transformations, rather than functions between sets. A mathematical situation is a model of a theory if the functions that describe it obey the equations of motion of the theory.

Passive theoretical symmetry: A transformation f from a model m of a theory Φ to a mathematical situation $f(m)$, is a *passive theoretical symmetry*

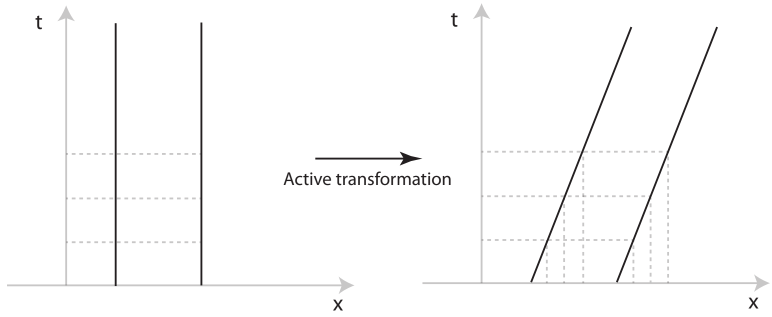

Figure 1.2: Representation of an active transformation in Galilean spacetime, where the worldlines could represent the front and back of Galileo's ship. Before the transformation the ship is at rest according to the x coordinate, whereas afterwards it is in motion according to this same coordinate.

if $f(m)$ is a model of Φ and m and $f(m)$ are different representations of *the same* physical situation.

Active theoretical symmetry: A transformation f from a model m of a theory Φ to a mathematical situation $f(m)$, is an *active theoretical symmetry* if $f(m)$ is a model of Φ and m and $f(m)$ are representations of *different* physical situations.

To make contact with our earlier discussion of empirical symmetries, we might say that if $\mathbf{x}(t)$ represents the position of Galileo's boat in the harbour, then $f(\mathbf{x}(t))$ represents the boat sailing away from the harbour at a uniform velocity v_x. In this case, we say that the active theoretical symmetry corresponds to an empirical symmetry.

In the example given above, Healey and Brading and Brown agreed that f was a symmetry of Newtonian mechanics. Furthermore, both the active and the passive interpretations of f produced symmetries. However, if we look very carefully at the way Healey defines the notion of symmetry, we see that some transformations will satisfy his definition without satisfying Brading and Brown's. Consider a different transformation, f' which takes the form

$$\mathbf{x}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} \to f'(\mathbf{x}(t)) = \mathbf{x}''(t) = \begin{pmatrix} x''(t) \\ y''(t) \\ z''(t) \end{pmatrix} = \begin{pmatrix} x(t) + v_x(t)t \\ y(t) \\ z(t) \end{pmatrix} \quad (1.5)$$

where $v_x(t)$ is now a function of time. Plugging $f'(\mathbf{x})$ into (1.3) for the x component gives

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(x + v_x(t)t) = \frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + \frac{\mathrm{d}^2}{\mathrm{d}t^2}(v_x(t)t) = \frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + \frac{\mathrm{d}^2 v_x(t)}{\mathrm{d}t^2}t + 2\frac{\mathrm{d}v_x}{\mathrm{d}t} \tag{1.6}$$

Because $v_x(t)$ is now a function of time, the extra terms remain and the equation does not reduce to the form (1.3). Thus, according to Brading and Brown's definition, f' is not a symmetry transformation. However, we still have the possibility of interpreting f' as a passive transformation, in other words, as a "strange" description of the same physical situation. In this case the coordinate system we are using is in a varying state of motion with respect to the original coordinate system. Healey's definition of symmetry makes no reference to the form of the equations that must be used to judge whether a certain mathematical situation is a model of a certain theory. Thus, we are free to use a different equation of motion to judge whether $f'(\mathbf{x})$ is a model of Φ. Since we have interpreted f' passively, we can view the steps of the calculation (1.6) as a derivation of the laws of motion that a mathematical situation must obey to be a model of Φ from that particular coordinate system. Thus, we conclude that, according to Healey's definition, $f'(\mathbf{x})$ is a theoretical symmetry.

On the other hand, if we take an active interpretation of f', then we do not have the possibility of changing the form of the equations of motion. This is because f' represents the motion of a different particle, as described in the same coordinate system in which the laws hold in their original form (1.3). In this case, the steps of the calculation (1.6), when multiplied by the mass of the particle, represent a derivation of the forces that would need to be applied to the particle to make it follow the trajectory given by $f'(\mathbf{x})$. Since a force is necessary to make the particle follow this trajectory, it cannot be a model of Newtonian mechanics with no forces acting on the particle.

The example just discussed brought to light two important considerations involved in deciding whether a given theoretical transformation should be considered a symmetry of a theory. The first is that the *form* of the equations of motion which are used to judge whether a given mathematical situation is a model of a theory must be carefully specified in any definition of symmetry. The second is that the interpretation of a theoretical transformation as active or passive can affect whether the transformation is a symmetry or not. In chapter 5, we will discuss this issue at length. We will show that in some theories passive and active transformations are logically equivalent, but that this equivalence breaks down in other theories.

Chapter 2

Gauge transformations: the received view

In the first chapter, I introduced the distinction between empirical and theoretical symmetries, as well as that between active and passive transformations. I presented some preliminary criteria for a transformation to count as a symmetry, and highlighted some difficulties that any attempt to define the notion of symmetry must address. I illustrated the analysis with examples from Newtonian mechanics.

In this chapter, I will present Wigner's philosophy of symmetry, which is the canonical interpretation of symmetry in modern physics. I will also introduce the notions of gauge transformations, and gauge symmetry (also referred to as gauge invariance). We will see what place Wigner gives these transformations within his analysis of symmetry. Because Wigner's views have been so influential, later contributions to the literature are best appreciated in their light. At the end of this chapter, I will mention some ways in which the current literature on symmetry builds on Wigner's work.

2.1 Wigner's classification of symmetries

In 1984, in the opening paragraph of *The Meaning of Symmetry*, Wigner states [Wigner, 1992a]:

> We are fundamentally discussing a question of language: what we now call symmetry. But questions of language are not unimportant. After all, we want to communicate with each other. In physics, the word symmetry has been used, in my opinion, in three different senses. I will try to describe them and tell you

why I do not like the idea of gauge invariance being a symmetry principle.

In the meantime, questions about "the meaning of symmetry" have developed into a foundational concern that is quite alot more than just "a question of language". The empirical success of physical theories that rely explicitly on symmetry principles (from special relativity to gauge theories, via general relativity) have made it imperative to fully understand the nature of symmetry in general, and of gauge symmetry in particular. Moreover, understanding the relation of gauge transformations to spacetime transformations is crucial if we are to make progress towards unifying our fundamental physical theories. Before discussing Wigner's philosophy of symmetry in more detail, I introduce the notions of a gauge transformation and gauge symmetry.

2.1.1 A brief introduction to gauge transformations and gauge invariance

Gauge transformations take place in the mathematical domain (as defined in chapter 1), and can be expressed, in accordance with Brading and Brown's definition, in terms of transformations of the dependent and independent variables of a theory describing electromagnetic processes. The precise definition of a gauge transformation depends on whether it acts on a classical or a quantum system.

In a non-relativistic formulation of classical electromagnetism, gauge transformations act jointly on the electric scalar potential ϕ and the magnetic vector potential $\mathbf{A}$ as

$$\begin{cases} \phi \to \phi' = \phi - \frac{\partial \chi(\mathbf{x},t)}{\partial t} \\ \mathbf{A} \to \mathbf{A}' = \mathbf{A} + \nabla \chi(\mathbf{x}, t) \end{cases} \tag{2.1}$$

with $\chi(\mathbf{x}, t)$ a smooth scalar function of space and time. If $\chi(\mathbf{x}, t)$ is a constant, the transformation is usually called global, otherwise it is called local. However, this way of making the global/local distinction in notoriously problematic, and we will question it in chapter 5. In a relativistic formulation, gauge transformations can be expressed in terms of the electromagnetic 4-potential $A^\mu = (\phi, \mathbf{A})$

$$A^\mu \to A'^\mu = A^\mu - \partial^\mu \chi(\mathbf{x}, t) \tag{2.2}$$

Gauge symmetry (or invariance) refers to the fact that the electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$ are invariant under a gauge transformation of the potentials. In a non-relativistic setting, this can be shown by looking at how the **E** and **B** fields transform under gauge transformations.

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{2.3}$$

becomes

$$\mathbf{B}' = \nabla \times (\mathbf{A} + \nabla\chi(\mathbf{x}, t)) \quad = \nabla \times \mathbf{A} + \nabla \times \nabla\chi(\mathbf{x}, t)) = \nabla \times \mathbf{A} = \mathbf{B} \tag{2.4}$$

because the curl of a gradient vanishes. For the electric field,

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} \tag{2.5}$$

becomes

$$\mathbf{E}' = -\nabla(\phi - \frac{\partial\chi(\mathbf{x}, t)}{\partial t}) - \frac{\partial}{\partial t}(\mathbf{A} + \nabla\chi(\mathbf{x}, t)) = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} = \mathbf{E} \tag{2.6}$$

because partial derivatives commute. In relativistic notation, the invariance can be shown more straightforwardly in terms of the electromagnetic field tensor

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{2.7}$$

which transforms as

$$F'^{\mu\nu} = \partial^\mu(A^\nu - \partial^\nu\chi(\mathbf{x}, t)) - \partial^\nu(A^\mu - \partial^\mu\chi(\mathbf{x}, t)) \quad = \partial^\mu A^\nu - \partial^\nu A^\mu = F^{\mu\nu} \tag{2.8}$$

In classical electromagnetism, all physically relevant quantities are taken to be functions of the **E** and **B** fields. Gauge transformations do not therefore result in any physical change. This suggests that they can be interpreted passively, as changes in the description of electromagnetic situations. In chapter 5 we will discuss at length whether they can also be interpreted actively.

When acting on a quantum system, described by a wavefunction $\psi(\mathbf{x}, t)$, possibly in the presence of an electromagnetic field, gauge transformations are *joint* transformations of the electromagnetic potential and the wavefunction, given by [Wigner, 1970, p. 23].

$$\begin{cases} \psi(\mathbf{x},t) \to \psi'(\mathbf{x},t) = e^{iq\chi(\mathbf{x},t)}\psi(\mathbf{x},t) \\ \phi \to \phi' = \phi - \frac{\partial\chi(\mathbf{x},t)}{\partial t} \\ \mathbf{A} \to \mathbf{A}' = \mathbf{A} + \nabla\chi(\mathbf{x},t) \end{cases} \tag{2.9}$$

This transformation also results in no change in the physical state of the quantum system. However, this is slightly more difficult to show than in the classical case, and details are given in chapter 4. In the context of a classical complex scalar field $\theta(x)$, or a Dirac field $\Psi(x)$ (where the dependence is now on a spacetime position represented by a 4-vector x), gauge transformations are respectively

$$\begin{cases} \theta(x) \to \theta'(x) = e^{iq\chi(x)}\theta(x) \\ A^\mu \to A'^\mu = A^\mu - \partial^\mu\chi(\mathbf{x},t) \end{cases} \tag{2.10}$$

and

$$\begin{cases} \Psi(x) \to \Psi'(x) = e^{iq\chi(x)}\Psi(x) \\ \bar{\Psi}(x) \to \bar{\Psi}'(x) = e^{-iq\chi(x)}\bar{\Psi} \\ A^\mu \to A'^\mu = A^\mu - \partial^\mu\chi(\mathbf{x},t) \end{cases} \tag{2.11}$$

These transformations leave the Lagrangians of their respective theories invariant, and therefore also do not change the physical states of the fields they act on (more details in chapter 4).

In our future discussions of gauge transformations and gauge symmetry, it should be clear from the context which sense of gauge transformation we are referring to. Otherwise, we mean the general fact that such transformations can be made whatever variant of an electromagnetic situation one chooses.

2.1.2 Wigner's philosophy of symmetry

In the passage quoted at the start of this chapter, Wigner claims to distinguish three different senses of the word symmetry. These are (1) symmetries of objects; (2) symmetries of the laws of nature; (3) gauge symmetries. However, he believes that this third sense is inappropriate, and that gauge symmetries have a different status from the other two. Later in the *The meaning of symmetry*, Yang suggests to Wigner that the reason for this difference is the impossibility of constructing two experiments which are gauge transforms of one another [Wigner, 1992a, p. 365]. Ultimately, it seems that both Yang and Wigner believe that gauge transformations can only be interpreted passively as changes in the description of a physical system. Wigner sums this up by claiming that gauge invariance "does not express anything physical". Yang agrees, and adds that gauge invariance should be seen as a new kind of symmetry, namely one that "tells us how interactions are formed", or in slogan form "symmetry dictates interactions".

This short exchange between Wigner and Yang highlights three issues still addressed today in the foundations of gauge theory. These are: (1) do gauge transformations have a dual interpretation in terms of active and passive versions? (2) what is the relation between gauge transformations and spacetime transformations? (3) how does gauge symmetry play a role in specifying the mathematical form of an interacting theory? The third question is a reference to the "gauge principle", which will be discussed in chapter 4. The first two questions will be the subject of chapter 5. In the rest of this section I will present Wigner's philosophy of symmetry, and clarify why he sets the notion of "gauge symmetry" apart from the others.

For Wigner, the world is complicated but the scientist, ideally, would like to find some order in this chaos. This can be achieved, first by separating out initial conditions from laws of nature, and secondly by recognizing that the laws of nature satisfy certain invariances [Wigner, 1970, p. 3]. The first step, while to some extent arbitrary, allows the scientist to focus on only a limited subset of the properties of the situation under study, and thus makes noticing correlations between these properties a tractable problem. The regularities in these correlations are the laws of nature. However, these regularities could never be formulated if the correlations observed by the scientist changed from day to day. Furthermore, the results of a scientist working here would have no relation to those of one working there if the laws of nature changed from place to place. Therefore, the fact that the laws of nature obey certain invariances is crucial for the possibility of their discovery.

Our knowledge of the world can therefore be seen to progress up a hierarchy, from the chaotic events that surround us, through the laws of nature according to which we order them, to the invariance principles which make the discovery of the laws of nature possible. In Wigner's philosophy, the invariances satisfied by the laws of nature thus arguably play the role of Kantian "transcendental principles" [Brading and Castellani, 2003, p. 1361] and [Mainzer, 1988]. More precisely, Wigner believes that there are four invariance (or symmetry) principles that the laws of physics obey: (1) position invariance; (2) time invariance; (3) rotation invariance and (4) uniform motion invariance. The first three he claims are fairly evident, but the discovery of the importance of the fourth he attributes to Einstein (although, as we have seen, Galileo also recognized this invariance) [Wigner, 1991, p. 197]. To say that the laws of nature "obey" these principles means that if a system is submitted to the relevant transformation (its position is changed, it is observed at a different point in time, it is rotated around some axis of space or it is set into a state of uniform motion), then it will continue to evolve according to the same laws as it did before the transformation. This insensitivity of the evolution of the system to the action of certain transformations is what Wigner means when he says that the invariance principles "express something physical". In this context, "physical" for Wigner means that something is actually done to the system to change it in some way. However, because he believes that gauge transformations cannot be interpreted as actually doing something to the system, he does not think that they can play the same role with respect to the laws of nature as the four other transformations just mentioned.

This distinction between the invariances that "express something physical" and those that do not is formalized by Wigner as the distinction between *geometric* and *dynamical* transformations. He also refers to this distinction as that between the "old" and the "new" invariance principles. The new principles represent an extension of the concept of symmetry into "an area where its roots are much less close to direct experience and observation than in the classical area of spacetime symmetry" [Wigner, 1970, p.15]. Although Wigner does not see it this way, we should take this comment as a warning. We should be careful not to be misled by the fact that we are simply less familiar with the notion of gauge transformations from our everyday experience. If we were different creatures, with sense organs sensitive to subtle electromagnetic interference effects, then the notion of gauge transformations might be much more familiar to us.

Geometric invariance principles

The geometric invariance principles are those formed by the transformations of the Poincaré group, also known as the inhomogeneous Lorentz group. These are translations in space and time, rotations and Lorentz boosts. This last transformation corresponds to the setting of a system into a state of uniform motion, and in this respect it does not differ essentially from the invariance principle already expressed by Galileo. For Wigner, Einstein's important contribution was to re-express the notion of Galilean relativity and make it compatible with the development of electromagnetic theory. He stresses that these geometric principles, as well as structuring the laws of nature, also can be formulated "in terms of the events themselves" [1][Wigner,

[1] This statement can be given a precise mathematical formulation as a point transformation, see chapter 5.

1970, p.17].

The notion of geometric invariance can be given the following (semi)formal treatment. Consider a system in a state ϕ. Now apply a transformation P_α to it, with P_α a member of the Poincaré group, such that the new state after the transformation is $\phi_\alpha = P_\alpha\phi$. ϕ and ϕ_α represent two physically different states. For instance, P_α might change the position of the system, or set it into a state of uniform motion. In a time t, the evolution of the system is given by the transformation T_t, with T representing some arbitrary dynamical law. In a time t, ϕ evolves into the state $T_t\phi$, and the state ϕ_α evolves into the state $T_t\phi_\alpha$. The fact that P_α is an invariance principle can be captured by the statement that $P_\alpha T_t\phi = T_t\phi_\alpha = T_t P_\alpha\phi$, in other words that the transformation and the dynamical evolution commute

$$[P_\alpha, T] = 0 \tag{2.12}$$

The significance of the commutation relation obeyed by P_α and T can be further elucidated by the following example, illustrated in figure 2.1. Imagine that P_α represents a change in position of the system, $x \to x'$, and that the dynamical evolution T is not invariant under such a transformation. In this case, T is written $T(x)$ (to indicate it has a position dependence) and $P_\alpha T_t(x)\phi \neq T_t(x)P_\alpha\phi$ because after performing the transformation, the system will be in a new position x' and, in general, $T(x) \neq T(x')$ when $x \neq x'$.

Dynamical invariance principles

A dynamical principle of invariance is one that "tells us how interactions are formed", as Yang put it. Gauge invariance, understood as the joint transformation of a matter field and a potential is an example thereof (details in chapter 4). Interestingly, Wigner believes that the general covariance of general relativity should also be considered as a dynamic principle of invariance [Wigner, 1970, p. 23]. He justifies this claim with a reference to Utiyama's work. Utiyama was one of the first to attempt (and partially succeed) to derive the equations of general relativity, for instance the form of the covariant derivative, by demanding the invariance of an arbitrary set of fields under "generalized Lorentz transformations" [Utiyama, 1956]. These are obtained by replacing the six parameters of the usual Lorentz group (three parameters for spatial rotations and three for boosts) with arbitrary functions of space and time. The usual Lorentz transformations, as well as those of the Poincaré group, are often referred to as "global" transformations. As in the case of gauge transformations, this is because they depend on a finite number of parameters rather than functions of space

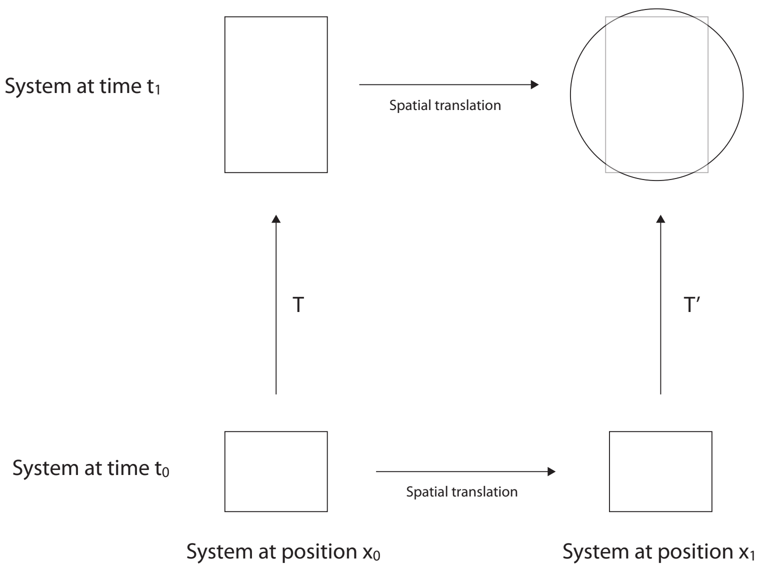

Figure 2.1: Example of the non-commutation of the evolution and transformation operators. Here P is a spatial translation. Because T depends on position, the result of evolving and then transforming (a rectangle) is different from the result of transforming and then evolving (a circle).

and time. When these parameters are changed into arbitrary functions of space and time, as in the case of the "generalized Lorentz transformations" employed by Utiyama, the transformations become "local".

The crucial point about dynamical principles of invariance is that they are based around *passive* transformations, ones that only change the descriptions of physical systems. For this reason, they cannot be used in the same way as the geometric invariances to infer something about the stability of the laws of nature across change. For such physical knowledge to be inferred from an invariance principle, the system must be physically transformed, something a passive transformation does not do.

As we showed in section 2.1.1, when the electromagnetic potential A_μ that describes a certain electromagnetic situation undergoes a gauge transformation of the type (2.2), the physical situation described by the new potential A'_μ is identical to that before the transformation. Thus the very same physical situation (understood as the same configuration of electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$), can be represented by two different electromagnetic potentials, A_μ and A'_μ. For this reason, Wigner compares A_μ to a "ghost", and the gauge transformations to "changes in the coordinates of that ghost" [Wigner, 1970, p. 22]. Given that ghosts, for Wigner, are invisible and powerless to affect the world they glide through, it seems that A_μ should be an equally impotent creature. However, it plays a crucial role in the dynamical invariance principle that determines the form of the equations describing the interaction of a quantum particle (or field) with an electromagnetic field. Troubled by this, Wigner admits to having tried to formulate these equations (in particular the Dirac equation) without reference to A_μ but was unsuccessful [Wigner, 1992a, p. 365]. Others have also tried to achieve this, but so far no one has succeeded [Zee, 2010]. Wigner admits that it is a "serious matter" that a "ghost" such as the electromagnetic potential should play such an important role in the equations of the electromagnetic interaction. This tension lies at the heart of modern attempts to interpret gauge symmetry.

Summary of Wigner's views

Wigner's philosophy presents a fragmented view of symmetry in modern physics. The Poincaré transformations acquire a special status as geometrical symmetries. This status derives from the possibility of interpreting them actively as changing the physical state of a system. Sharply distinguished from them are the dynamical symmetries, which include general coordinate transformations and gauge transformations. These have *only* a passive interpretation, and therefore cannot have the same physical significance as the geometric symmetries. Nevertheless, Wigner does believe that both general coordinate transformations and gauge transformations play similar roles in their respective theories by determining the form of the interaction equations. In the latter case, he expresses discomfort about the important role of the electromagnetic potential A_μ, which he compares to a "ghost". He does not show similar misgivings about the Christoffel symbols, which we will argue are the analogue of A_μ in GR. Note that all the transformations that Wigner discusses can be understood as occurring in the mathematical domain, in other words, they are transformations which act on mathematical structures.

2.1.3 After Wigner

The modern literature on symmetry still respects many of the basic tenets of Wigner's analysis. The sharp division between the status of the Poincaré transformations and the others, as well as the reason therefor, is echoed by Brown and Brading [Brading and Brown, 2004, p. 663]

> Global spacetime symmetries [those corresponding to the Poincaré transformations] have a special status, both theoretically and practically: theoretically they have an active interpretation in the sense that a symmetry transformation applied to a subsystem of the universe yields an empirically distinct scenario; and, furthermore, instances of these active transformations are implementable in practice through the use of effectively isolated subsystems. Neither global internal symmetries [global gauge transformations], nor local symmetries of either variety [general coordinate transformations and local gauge transformations], have even a theoretical active interpretation of this kind.

Wigner was also not the only one to take seriously Utiyama's attempts to formalize the analogy between general coordinate transformations and gauge transformations. An entire research program in theoretical physics has devoted itself to formulating general relativity as a gauge theory constructed in terms of the local transformations of a particular gauge group. Kibble and Sciama rapidly followed in Utiyama's footsteps [Kibble, 1961], [Sciama, 1964]. Later contributors to this tradition are, among others, Trautman, Ivanenko, Sardanashvily and more recently, Gronwald and Hehl [Trautman, 1980], [Ivanenko and Sardanashvily, 1983], [Gronwald and Hehl, 1996].

Wigner's worry about the "ghostly" A_μ gauge field playing an important role in the physics has been formalized by Redhead in terms of the relation between the "surplus mathematical structure" of a theory and its physical content. There have been several efforts to solve the problem by giving formulations of gauge theories that do not exhibit the gauge freedom of the electromagnetic potential. For instance, some advocate a formulation of electromagnetism based on holonomies (line integrals around loops in spacetime), which are gauge invariant quantities [Healey, 2007], [Belot, 1998, p. 543]. Similarly, DeWitt has argued for a formulation of quantum theory that does not rely on the electromagnetic potential [DeWitt, 1962].

Although there seem to be no explicit rejections of Wigner's philosophy in the literature, some recent developments, especially among philosophers of physics, signal the appearance of new attitudes. While the "symmetry dictates interaction" paradigm remains strong in the physics community (see Weinberg and 't Hooft), dissident voices are dominant among philosophers, who are keen to downplay its physical significance [Brown, 1999], [Martin, 2002], [Redhead, 2003]. As for the relation between gauge symmetries and spacetime ones, Martin has claimed that "in assessing the physical content of gauge symmetry principles, any analogy with the 'conceptual foundations' of GR (or of spacetime theories generally) is perhaps more trouble than it is worth" [Martin, 2003, p. 57]. This is perhaps the clearest break yet with Wigner's legacy. It may well have its origins in Earman's attempts to promote an analysis of "gauge" in GR, and in other theories, by means of the constrained Hamiltonian formalism [Earman, 2002]. In this framework, the notion of "gauge" takes on a different significance from the geometrical one which I will defend in this thesis.

In the last sections of this chapter I will review two recent contributions to the philosophy of gauge theory which have clear roots in Wigner. The first is Redhead's notion of surplus structure, and the second is Kosso's notion of "observing a symmetry", further elaborated by Brown and Brading [Kosso, 2000], [Brading and Brown, 2004]. I will also relate our considerations in chapter 1 to Wigner's views.

The concept of surplus structure

Wigner's discomfort about the role of the electromagnetic potential is captured by Martin's identification of a tension between "the redundancy of gauge and the profundity of gauge" [Martin, 2003, p. 52]. Redhead's analysis of this tension in terms of the concept of "surplus structure" formalizes the issue [Redhead, 2003], [Redhead, 2001], [Redhead, 1975]. It takes place in the context of the approach to the philosophy of science that focuses on representation and appeals to the mathematical tools of model theory. The starting point is the "empirical-historical fact that theories in physics can be represented as mathematical structures" [Redhead, 1975, p. 87]. The notion of representation is formalized according to the model theoretic concepts of structures, embeddings and isomorphisms. The physical world is represented as a structure P which is then embedded in a larger mathematical structure M' which is itself a representation of a certain theory applicable to the domain of P. This embedding is defined by means of an isomorphism between P and a substructure M of M' [Redhead, 2003, p. 126].[2] M' is larger than M (it contains elements and relations that are not included in M) and the relative complement of M in M' is what Redhead calls "surplus structure" [Redhead, 2003, p. 128] .

The key feature of surplus structure is that is has no "physical correlate", in other words it does not correspond to anything in the world, as represented by P. Redhead admits that as science progresses, what was once considered surplus structure can be found to actually have a physical interpretation. He mentions molecules in the time of Ostwald and Mach and energy in nineteen-century physics as examples [Redhead, 1975, p. 88],[Red-

[2] Note that one of the big problems with this approach is forming the set P in the first place, because this involves carving up the world in a certain way. This carving up of the world risks appearing hopelessly arbitrary and threatens to undermine any philosophical insight the "representation" approach to the philosophy of science promises to offer. For more details see van Fraassen's book *Scientific Representation*.

head, 2003, p. 129]. However, this does not always have to be the case. He claims that the S-matrix theory of elementary particles never intended to associate anything physical with the surplus structure. Questions of historical evolution aside, the most remarkable feature of surplus structure is that it can still be useful for making inferences about elements in P, even if it seems to have no direct correspondence to anything in P. As an example, Redhead mentions the use of complex currents in alternating current theory. Clearly, a current in the world cannot be complex valued, but by representing it in this way, calculations can be made which were not possible or not so easy otherwise. Complex numbers can also be used in classical mechanics to represent oscillations. Once again, the position of a particle cannot be complex valued, but calculations can be simplified by using the exponential notation for complex numbers, and the solutions thus obtained can then be converted back into sensible results (for instance by taking the real part).

The notion of surplus structure should not be confused with that of "theoretical terms" or "unobservable entities", although the two may have some areas of overlap. One major difference is that surplus structure is meant also to accommodate strictly mathematical methods and objects which would not fit neatly into an account in terms of theoretical terms. Complex valued currents are not "unobservable", even in principle, yet they still play a role in deriving empirically relevant predictions. It is this role that the concept of "surplus structure" is meant to account for.

Redhead suggests that gauge transformations should be understood as "automorphisms of M' that reduce to the identity on M" [Redhead, 2003, p. 129]. This means that a gauge transformation does something to the surplus structure but not to the rest. In this context, Wigner's ghost takes the form of the surplus structure, and the mystery becomes why properties of this surplus structure should have consequences for the substructure M and therefore for P. Unfortunately, the precise way in which the properties of the surplus structure affect M is left vague, and Redhead is forced into making statements of the kind "physical structure being controlled by requirements imposed on surplus mathematical structure" [Redhead, 2003, p. 131]. Elucidating the precise nature of this "control" thus becomes a primary concern of someone who subscribes to Redhead's analysis.

Another approach to surplus structure is to consider it as a kind of "fat" that should be trimmed away from the theory proper, leaving only the substructure M' which has well defined physical correlates. In the case of gauge transformations as surplus structure, such a view leads to formulating a theory only in terms of so-called "gauge invariants". These are the quantities that remain unchanged under gauge transformations. However, repeated failure to eliminate the electromagnetic potential from quantum-electrodynamics suggests that such an approach is not actually feasible.

This deepens the mystery: not only is the surplus structure useful, but it seems to be essential as well. Martin and Redhead both acknowledge that such considerations seem to point towards a quasi-Platonic interpretation of the role of mathematics in physics [Martin, 2003, p. 52]. Wigner has also expressed his amazement at the success of mathematical methods in physics [Wigner, 1960].

Physical significance of gauge transformations

As stated at the end of section 2.1.2, Wigner believes that gauge transformations relate descriptions of the same physical system, and therefore that no physical operation on a system can implement a gauge transformation. This has the immediate consequence that gauge symmetry, the invariance of a system under a gauge transformation, can also not "express anything physical". On the other hand, we have also seen that Yang believes that gauge symmetry can be used to infer the form of the electromagnetic interaction. Surely, this should count as "expressing something physical"? A closer look at Wigner's philosophy of symmetry reveals that when he talks of physical significance, he has something very specific in mind. For Wigner, symmetries (or invariances) express something physical when they allow conclusions to be drawn about the laws of nature themselves. In this case, they warrant the claim that the laws of nature are the same at all times t, at all positions x, in all rotated states θ and in all states of uniform motion v. In the formalism of section 2.1.2, they state that the dynamical evolution operator T does not depend on the variables (t, x, θ, v).

One might be tempted to respond that gauge invariance entails that T also does not depend on a variable that would represent the gauge. But Wigner would not accept this line of reasoning. He would argue that it is trivial that T does not depend on the gauge, because changing the gauge does not change the system "in-itself", but merely its description. And if the system is not changed (if it is not physically interacted with), then it should be obvious that its evolution will remain unchanged!

Our considerations in chapter 1 show that this reasoning, no matter how evident it may seem, is not entirely satisfactory. Changing the way a system is described may also require a new way of describing its evolution (as we saw for the trajectory of a particle in Newtonian mechanics if described in a non-inertial coordinate system). However, this new evolution would only be apparently different from the old one, in the sense that the evolution of the system "in-itself" would not be changed. In this case, it seems that Wigner assumes that the evolution operator T is somehow "objective" in the sense that it describes the evolution of the system "in-itself". In contrast, doing something physical to a system, for instance changing its position, could feasibly have consequences for how it evolves “in-itself”. For these reasons, Wigner believes that invariance under a physical change is an important property of the laws of nature, whereas invariance under a change of description should not be.

One might take issue in this analysis with the use of the concepts of “system in-itself” and “evolution in-itself”. Despite the obvious Kantian connotations of these expressions, I believe that they are quite harmless. The “in-itself” is simply meant to capture the idea that physical systems in the world go about their usual business whether or not someone is describing them. By description, I mean something like the action of a painter, who paints a landscape without interfering with it. The notion of description is thus entirely distinct from the notion of measurement.

Wigner’s concerns about the physical significance of symmetries can be explicated using Kosso’s notion of “observing a symmetry” [Kosso, 2000, p. 82]. If a symmetry can be observed (in Kosso’s sense), then Wigner would agree that it “expresses something physical”. The two conditions for a symmetry to be observed are (as neatly summarized by Brading and Brown [Brading and Brown, 2004, p. 646]:

Transformation condition: the transformation of a subsystem of the universe with respect to a reference system must yield an empirically distinguishable scenario.

Symmetry condition: the internal evolution of the untransformed and transformed subsystems must be empirically indistinguishable.

These conditions should be familiar from chapter 1, although they are formulated here in a way that relies essentially on the notion of subsystem. Brading and Brown add this notion to Kosso’s account in order to make explicit the fact that it must be possible to determine that a transformation has taken place. If a given transformation is a symmetry of a system, then by definition performing this transformation does not change the evolution of the system. However, if we are to consider the transformation as having changed the system in some way, then something about the system must be different after the transformation. Applying transformations to subsystems of the universe allows for changes in the relations between these subsystems and the untransformed parts of the universe to act as witnesses for the transformation.

For the case of Wigner’s geometric transformations, it is easy to see that they all pass these two conditions, and thus qualify as observable symmetries. The cases of translations in space and time and rotations are straightforward. A Galileo ship type situation can serve to demonstrate symmetry under changes in the uniform motion of a system. On the other hand, the extension of this analysis to the dynamical symmetries is more problematic.

As far as Wigner is concerned, the dynamical symmetries would fail the transformation condition by definition: spacetime coordinate transformations and gauge transformations (the two common examples of dynamical transformations that Wigner gives) relate descriptions of a physical system, so it is obvious that performing such transformations doesn't yield an empirically distinguishable situation.

Conclusions

At the end of chapter 1, we saw that spactime transformations can have a dual interpretation in terms of active and passive versions. However, according to Wigner, this is not the case for gauge transformations. We might ask where this difference originates. Is it due to a fundamental fact about the nature of gauge transformations or is it simply the artifact of Wigner's particular interpretational approach? We also saw that Wigner thought that general coordinate transformations and gauge transformations played similar methodological roles in GR and Q-EM respectively. The methodological role of passive transformations will be the subject of chapter 4. My conclusions will support Wigner's views. However, in chapter 5 I will argue that gauge transformations can be interpreted actively, and this will break with Wigner's legacy.

Chapter 3

Mathematical notions

In the first two chapters I discussed the status and role of transformations and symmetries in physics while relying to a large extent on the reader's background knowledge and physical intuition. In the next three chapters, I will address the same issues, but set them in a well defined mathematical setting. Apart from our treatment of non-relativistic quantum mechanics, the tools will be primarily geometrical.[1]

As I explained in chapter 1, transformations in physical theories are defined in the mathematical domain. This means that they act on, and between, mathematical structures. Since a mathematical structure is a representation of a physical situation, an interpretation of the structure is needed to understand what exactly it represents. Different structures may represent different systems, but they may also represent the same system in different ways. As Fonda and Ghirardi observe, a good theory should do more than just give a mathematical description of certain physical situations [Fonda and Ghirardi, 1970, p. 11]:

> The theory of a physical system, in fact, is fully defined only if it also contains the specification of the connection existing between the descriptions of all possible states of the system when viewed by different observers.

In other words, the theory must also have something to say about how different representations of the same system (perhaps corresponding to the descriptions of different observers) relate to each other. In this section, I introduce some of the technical machinery needed to achieve this in the context of spacetime theories, non-relativistic quantum mechanics and gauge

[1]Perhaps some would argue that the mathematics of Hilbert spaces is geometry. However, since there are no notions of points, or curves in Hilbert space, I do not consider this setting geometrical in the sense of differential geometry.

theory.

3.1 Spacetime notions

In this thesis, I will mostly operate within a "coordinate-based" approach to spacetime theories. Since one of my major concerns is elucidating how the freedom that exists in describing a system can be exploited in formulating physical theories about the world, it makes sense to take an approach in which this freedom is clearly present.[2]

3.1.1 Manifolds

Manifolds are the setting of spacetime theories. I provide the necessary definitions before discussing their interpretation. The definition of a manifold requires those of a topological space, a homeomorphism and a neighbourhood.

Definition 1. [Dieks, 2011, p. 108] A *topological space* is a pair $\{\mathcal{X}, \mathcal{T}\}$ where $\mathcal{X}$ is a set and $\mathcal{T}$ is a family of subsets of $\mathcal{X}$ such that

1. for any collection of members of $\mathcal{T}$ their union is in $\mathcal{T}$;
2. for any finite collection of members of $\mathcal{T}$ their intersection is in $\mathcal{T}$;
3. $\mathcal{X}$ and $\emptyset$ are in $\mathcal{T}$.

Definition 2. [Isham, 1999, p. 51] A *homeomorphism* is map $f : \{\mathcal{X}, \mathcal{T}\} \to \{\mathcal{Y}, \mathcal{T}'\}$ between topological spaces such that

1. f is a bijection
2. f and its inverse f^{-1} are continuous

The neighbourhood of a point can be understood intuitively as a set containing the point and points around it in a way that certain considerations of convergence are respected (more details in [Isham, 1999, p. 25]). This allows us to define a manifold.

Definition 3. [Torretti, 1983, p. 257] An n-dimensional manifold $\mathcal{M}$ is a topological space $\{\mathcal{X}, \mathcal{T}\}$ such that every point of $\mathcal{X}$ has a neighbourhood homeomorphic with $\mathbb{R}^n$.

[2] When discussion the fibre bundle formulation, the tools will become those of coordinate-free differential geometry. However, we will see that, despite this, a certain freedom in the description of a "gauge system" creeps back in.

The manifold $\mathcal{M}$ is often interpreted as the "arena" in which our physical world unfolds (Minkowski poetically called the 4-dimensional manifold in which he set special relativity "the world" [Torretti, 1983, p. 20]). Each point in the manifold (each element of the set $\mathcal{X}$), represents an *event*. Geroch defines an event as "an idealized occurrence in the physical world having extension in neither space nor time", such as "the explosion of a firecracker" or the "snapping of one's fingers" [Geroch, 1978, p. 3]. However, as Geroch himself notes, such a definition seems to entail that something must happen at a point of a manifold for it to "exist". But this is too restrictive for the needs of physics. In this way, Torretti suggests that we should consider points of the manifold as "possible events", and the manifold itself as the "collection of all punctual instantaneous locations available for them" [Torretti, 1983, p. 22]. This allows us to use the concept of a manifold to represent a universe that is mostly empty, without the mathematical structure collapsing onto only those points where something happens. As we will see later, the famous "hole argument" raises additional problems for the interpretation of the points of a manifold.

In order to be able to do physics, in other words to formulate laws relating the events represented in the manifold, a way of identifying, or naming, the events is needed. This is the role of coordinate systems.

3.1.2 Coordinate systems and coordinate transformations

In differential geometry, coordinate systems are often referred to as coordinate charts. From now on, in order to harmonize with practices in physics, I will use the term coordinate system rather than coordinate chart. Roughly, a coordinate system associates sets of n real numbers (called the coordinates) to points of the n-dimensional manifold. Due to the geometrical properties of the manifold, it may not be possible for one coordinate system to associate coordinates to each point. In this case, the manifold much be covered by coordinate "patches". This is the case for the points on the surface of a sphere for example, for which two coordinate systems are needed. [3] It is possible for these patches to overlap, in fact, there are many different ways of assigning coordinates to the points of a manifold. Transformations between these different ways of assigning coordinates are called "coordinate transformations". Now for the formal definitions:

Definition 4. [Isham, 1999, p. 61] An n-dimensional *coordinate system* on a manifold $\mathcal{M}$ is a pair (U, ϕ) where U is an open subset of $\mathcal{M}$ (called the domain of the coordinate system) and $\phi : U \to \mathbb{R}^n$ is a homeomorphism of U into an open subset of the Euclidean space $\mathbb{R}^n$. If $U = \mathcal{M}$, then the coor-

[3] This is related to the problem of not being able to comb a "hairy" sphere.

dinate system is said to be globally defined; otherwise it is locally defined.

Definition 5. [Isham, 1999, p. 63] A point p in an open subset U of $\mathcal{M}$ has coordinates $(\phi^0(p), \phi^1(p), \ldots, \phi^{n-1}(p)) \in \mathbb{R}^n$ with respect to the coordinate system (U, ϕ). The *coordinate functions* $\phi^\mu : U \to \mathbb{R}$, $\mu = 0, 1, \ldots, n-1$, are often written as x^μ, and the coordinates of a particular point p as $(x^0(p), x^1(p), \ldots, x^{n-1}(p))$.

In what follows, I will often refer to a coordinate system simply by means of its coordinate functions. Thus, rather than writing "in the coordinate system (U, ϕ)", I will write "in the coordinate system x^μ". For coordinate systems (U_1, ϕ_1) and (U_2, ϕ_2) which overlap, $U_1 \cap U_2 \neq \emptyset$, one can define a coordinate transformation from the coordinates of a point $p \in U_1 \cap U_2$ with respect to (U_1, ϕ_1) to those with respect to (U_2, ϕ_2).

Definition 6. [Isham, 1999, p. 61] Let (U_1, ϕ_1) and (U_2, ϕ_2) be a pair of n-dimensional coordinate systems with $U_1 \cap U_2 \neq \emptyset$. Then the coordinate transformation between the two coordinate systems is the map $\phi_2 \circ \phi_1^{-1}$ from the open subset $\phi_1(U_1 \cap U_2) \subset \mathbb{R}^n$ into the open subset $\phi_2(U_1 \cap U_2) \subset \mathbb{R}^n$. Thus a coordinate transformation $\phi_2 \circ \phi_1^{-1}$ is map from a subset of $\mathbb{R}^n$ to a subset of $\mathbb{R}^n$.

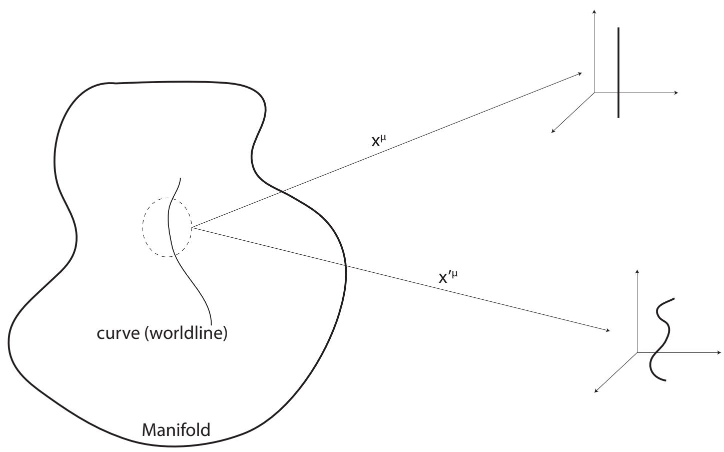

Figure 3.1: Illustration showing how the same section of a curve in the manifold is coordinatized in different ways by different coordinate systems x^μ and x'^μ.

3.1.3 Reference frames

The previous definitions establish the notion of a coordinate systems as a piece of mathematical machinery. However, physicists are interested in making measurements in the real world, and therefore some bridge is necessary from the distance measurements and clock readings in the laboratory to the n-tuples of real numbers assigned by coordinate systems to events. Up to now, we have considered manifolds of arbitrary dimension, but from now on we will mainly be concerned with 4-dimensional manifolds. This is no surprise, given that our spacetime descriptions usually take the form of a position measurement (involving three numbers) and a clock reading (involving one number).

A *reference frame* is the general name given to a collection of physical objects capable of making measurements (in spacetime theories, usually distance and time measurements). In special relativity for example, a reference frame consists of a set of rigid rods and clocks synchronized using Einstein's light beam procedure. When an event occurs, a 4-tuple of numbers (X^0, X^1, X^2, X^3) can be assigned to it by using the reference frame in the following way: X^0 indicates the reading of the clock at the location at which the event occurs, and X^1, X^2, X^3 count the number of rods in each of the three spatial dimensions that separate the location of the event from the origin of the reference frame. In this way, a reference frame, just as a coordinate system, can be used to assign 4-tuples of real numbers to spacetime events.

Despite the obvious similarities, there are important differences between reference frames and coordinate systems. The main source of the differences is that reference frames, being physical objects, are subject to all sorts of restrictions, whereas coordinate systems, being mathematical objects, are far more flexible (they are subject only to smoothness and invertibility requirements). This means that there are coordinate systems that assign 4-tuples of numbers that would not correspond to the measurements of any imaginable reference frame. In order to make the relationship between the two concepts clearer, it is useful to give a general idealization of a reference frame. The most common is to use a congruence of time-like worldlines (the time-like requirement comes from the fact that the constituents of the frame are physical objects, and therefore cannot move faster than light). This is a set F of time-like curves such that each point of the manifold lies on the range of one and only one curve of the set [Norton, 1993, p. 837],[Torretti, 1983, p. 28]. In general, we also require that the frame F be rigid, which means that the distance between any pair of curves in F stays constant throughout their history. In this way, each curve in F is taken to represent the worldline of a tiny piece of the physical reference frame.[4]

We can now give the precise relationship between a reference frame and a

[4]The fact that a reference frame can be idealized in such a way assumes that physical objects can be represented as systems of point particles in motion. This is a common assumption of spacetime theories.

coordinate system. Consider a physicist using a frame F to make measurements. Each curve in F will serve to fix the notion of "same point in time". This means that, in the 4-tuples of numbers assigned to each point along a curve in F, the three spatial values X^1, X^2, X^3 will stay constant, and only the time value X^0 will change. We say that a coordinate system (U, ϕ) is *adapted* to the frame F, if, for any two points p and p' along a curve W in F, $\phi^1(p) = \phi^1(p'), \phi^2(p) = \phi^2(p'), \phi^3(p) = \phi^3(p')$, and ϕ^0 is a monotonically increasing function up the curve that assigns times compatible with the Einstein synchronization procedure. In words, a reference frame and coordinate system adapted to it would assign the same state of motion to a particle represented by an arbitrary worldline in the manifold.

In the flat Minkowski spacetime of special relativity, inertial coordinate systems are equivalent to inertial reference frames. For this reason, it is easy to forget the differences between the two concepts. However, once one considers non-inertial coordinate systems (such as a rotating one for example), it becomes impossible to imagine a physical system that would assign 4-tuples of numbers in the same way as the coordinate system. As Norton remarks in the context of Einstein's struggles with the rotating disk, "he [Einstein] found the need to introduce coordinate times which could not be read directly from clock measurements" (Norton [1993], p.836). This problem becomes even more acute in the non-euclidean spacetime geometries of general relativity.[5] In GR, it becomes advantage to define reference frames by specifying, for each point of the manifold, one timelike and three orthonormal spacelike vectors. This is known as a frame field, or also as the vierbein formalism and it is useful in extracting physical predictions from general relativity.

3.2 Quantum notions

I will now present the mathematical tools needed to account for the relations between the descriptions of different observers in non-relativistic quantum mechanics.

In quantum mechanics, because all physical information that can be gained by making a measurement on a system is given by the transition probability

$$p(\phi, \psi) = \frac{|\langle\phi|\psi\rangle|^2}{\langle\phi|\phi\rangle\langle\psi|\psi\rangle} \tag{3.1}$$

and that $p(\phi, \psi)$ is invariant under either or both of the transformations

[5] In fact, the rotating coordinate system is a well studied case that showed the need for non-euclidean geometries even in the flat Minkowski spacetime of special relativity.

$$|\phi\rangle \to a|\phi\rangle$$
$$|\psi\rangle \to b|\psi\rangle$$

where a and b are arbitrary complex numbers, one generally accepts that the set of all $a|\phi\rangle$ (all $b|\psi\rangle$) represent the same physical state. This set is called a ray, written $\{\mathbf{\Phi}\}$ ($\{\mathbf{\Psi}\}$). The set of all unit vectors belonging to a given ray is called a unit ray [Fonda and Ghirardi, 1970, p. 5].

In spacetime theories, the most fundamental layer of reality is the manifold and the worldlines of particles therein. The next layer is the "coordinate view" (a description of the worldlines in terms of n-tuples of real numbers) that is assigned by a specific coordinate system. Coordinate transformations are then maps between n-tuples and they do not affect the worldline in the manifold. In the quantum mechanical formalism on the other hand, one could argue that there is nothing analogous to the manifold of the spacetime theories. Instead, the most fundamental possible description of a system is as the state vector of a Hilbert space. This state vector should not be interpreted as an abstract representation of the system in the absence of an observer (as one might consider the worldlines in the manifold) but instead as implicitly assuming a certain reference frame. A change of reference frame (corresponding to a coordinate transformation in a spacetime theory) is represented in quantum theory as a vector mapping (a mapping of the Hilbert space onto itself) which does change the state vector assigned to a particular system.

This correspondence between spacetime theories and quantum mechanics is sometimes made more confusing than necessary. For example, Auyang claims that "a basis of $\mathscr{H}$ is analogous to a coordinate system in Cartesian geometry" [Auyang, 1995, p. 19]. However, I will show that such a view is problematic and that is does not agree with the account given by Fonda and Ghirardi. First of all, I will explain why this view seems reasonable.

Definition 7. A *complete orthonormal basis* of an n-dimensional Hilbert space $\mathscr{H}$ is a set of n unit vectors $\{|\alpha_i\rangle\}$ such that any vector $|\phi\rangle \in \mathscr{H}$ can be written as a linear combination of the basis vectors, $|\phi\rangle = \sum_i^n c_i|\alpha_i\rangle$ with c_i complex numbers.

A Hilbert space $\mathscr{H}$ has an uncountable infinity of different bases. Furthermore, because a change of basis $\{|\alpha_i\rangle\} \to \{|\beta_j\rangle\}$ does not affect the state vector $|\phi\rangle$, but only its "description" in terms of basis vectors, it is tempting to see the state vector as analogous to a worldline, and a basis as analogous to a coordinate system. However, such an analogy does not capture how the theory actually functions to describe physical systems. Anticipating what will be formalized below, consider two observers O and $\bar{O}$, that describe the same quantum mechanical system S. O will describe S by a certain state vector $|\phi\rangle$, and $\bar{O}$ by a state vector $|\psi\rangle$. If O and $\bar{O}$ are using reference frames that are translated with respect to each other for example, then in general (in the Schrödinger picture), $|\phi\rangle \neq |\psi\rangle$. This shows that the analogy suggested by Auyang does not capture the way that transformations between the descriptions of different observers are represented in quantum mechanics. Two observers O and $\bar{O}$ are not related by a change of basis, and the state vector is not analogous to the worldlines of a system. Instead, each observer uses her own Hilbert space to describe a system, and thus each observer can choose to decompose the state vector she assigns to the system S in any number of different bases. In quantum mechanics, a change of reference frame thus corresponds to a vector mapping. I will now present a more formalized version of this story.

The problem at hand is how quantum mechanics deals with the fact that one system can be described by different observers. For now I will consider only two observers, O and $\bar{O}$. I assume that a reference frame can be associated to both observers. In the context of quantum mechanics, a reference frame simply means some set of instruments that can perform measurements on the system. Furthermore, I will assume that the reference frames are "macroscopic", meaning that they do not fall under the laws of quantum mechanics themselves. The implicit distinction thus introduced between the microscopic and the macroscopic world will not bother us at this stage. The reference frames associated with two different observers may be related in any number of ways, but here I will mostly assume that the transformations between the different reference frames are elements of the Poincaré group. This means that I consider translations in space and time, rotations in space and states of uniform motion. I will work within the Schrödinger picture and thus consider that the evolution of a system takes place in the state vector, and not in the operators. I will follow the approach of Fonda and Ghirardi [Fonda and Ghirardi, 1970, chapter 1].

Two observers O and $\bar{O}$, that differ in the way suggested above, describe the same quantum mechanical system S at a time t, by assigning to it a ray in Hilbert space. O assigns the ray $\{\mathbf{\Psi}_O(t)\}$ and $\bar{O}$ the ray $\{\mathbf{\Psi}_{\bar{O}}(t)\}$. It is then assumed that a one-to-one mapping $\boldsymbol{T}$ exists between the physically possible rays of O and those of $\bar{O}$, such that $\{\mathbf{\Psi}_{\bar{O}}(t)\} = \boldsymbol{T}\{\mathbf{\Psi}_O(t)\}$. The existence of such a mapping relies on the fact that the observations of O can in fact be translated into those of $\bar{O}$. This is similar to the requirement in the space-time case that the domains of two coordinate systems be overlapping for there to be a possible coordinate transformation between them. The states $|\phi_O\rangle \in \{\mathbf{\Phi}_O(t)\}, |\psi_O\rangle \in \{\mathbf{\Psi}_O(t)\}$ and $|\phi_{\bar{O}}\rangle \in \{\mathbf{\Phi}_{\bar{O}}(t)\}, |\psi_{\bar{O}}\rangle \in \{\mathbf{\Psi}_{\bar{O}}(t)\}$ assigned by O and $\bar{O}$ respectively to the system S are such that

$$|\langle\phi_{\bar{O}}(t)|\psi_{\bar{O}}(t)\rangle|^2 = |\langle\phi_O(t)|\psi_O(t)\rangle|^2 \tag{3.2}$$

This says that O and $\bar{O}$ must agree on the probability that after a measurement on the system in the state $|\psi_O(t)\rangle$ for O and $|\psi_{\bar{O}}(t)\rangle$ for $\bar{O}$, the system will be found in the state $|\phi_O\rangle$ for O and $|\phi_{\bar{O}}\rangle$ for $\bar{O}$. If this were not the case, but one still maintained that $\boldsymbol{T}$ was a translation from the "language" of O to the "language" of $\bar{O}$, then one would be forced to the very odd conclusion that the evolution of the system "in-itself" somehow depended on the way one chose to describe it. In other words, it would amount to the conclusion that changing the description of the system (which involves no interaction with the system) has consequences for the evolution of the system. This is ruled out *a priori*, and therefore we conclude that both O and $\bar{O}$ must agree on probabilities related in this way by the mapping $\boldsymbol{T}$.

We now wish to define a vector mapping between states of the Hilbert space, rather than rays. O and $\bar{O}$ describe S using the Hilbert space $\mathscr{H}$. A vector mapping T from $\mathscr{H}$ into itself is compatible with $\boldsymbol{T}$ if for every $|\psi\rangle \in \{\boldsymbol{\Psi}\}$ then $T|\psi\rangle \in \boldsymbol{T}\{\boldsymbol{\Psi}\}$. This means that T maps elements of the ray $\{\boldsymbol{\Psi}\}$ into elements of the ray $\boldsymbol{T}\{\boldsymbol{\Psi}\}$. Using this vector mapping T, (3.2) can be restated

$$|\langle T\phi_O(t)|T\psi_O(t)\rangle|^2 = |\langle\phi_O(t)|\psi_O(t)\rangle|^2 \tag{3.3}$$

I now state a famous theorem by Wigner which characterises this vector mapping T.

Theorem 1. *[Fonda and Ghirardi, 1970, p. 14] A surjective map $T : \mathscr{H} \to \mathscr{H}$ that satisfies $|\langle T\phi_O(t)|T\psi_O(t)\rangle| = |\langle\phi_O(t)|\psi_O(t)\rangle|$ for all $|\phi\rangle, |\psi\rangle \in \mathscr{H}$ has the form $T|\phi\rangle = aU|\phi\rangle$, where a is an arbitrary complex number of modulus 1 and U is a linear or anti-linear unitary operator.*

Since T satisfies (3.3) (the modulus of the inner product is always positive), it satisfies the conditions of Wigner's theorem, and therefore it must have the form of a linear or anti-linear unitary operator U.

These considerations can be summarized as follows. Consider two observers O and $\bar{O}$ that describe a quantum system S by ascribing to it states of a Hilbert space $\mathscr{H}$. If the transformation from the reference frame of O to that of $\bar{O}$ is an element of the Poincaré group, then the descriptions of O can be converted into the descriptions of $\bar{O}$ by the use of a linear or anti-linear (dependent on the relation between O and $\bar{O}$) unitary operator U (determined up to a phase factor). Thus, if O assigns to S the state $|\psi\rangle$, then $\bar{O}$ will assign it the state $U|\psi\rangle$.

Finally, a remark on the transformation properties of the operators used by O and $\bar{O}$ to make measurements on S. Imagine that the operator $\hat{q}$ represents the distance of the system from the center of the reference frame of O. In the reference frame of $\bar{O}$, this operator is written $\hat{q}' = U\hat{q}U^\dagger$, where † is the Hermitian conjugate. However, the operator $\hat{q}'$ will still give the distance of S from the center of the reference frame of O, because $\hat{q}'U|\phi\rangle = U\hat{q}U^\dagger U|\phi\rangle = U\hat{q}|\phi\rangle = qU|\phi\rangle$. If $\bar{O}$ wants to measure the distance of S from the center of her own reference frame, then she must also use the operator $\hat{q}$, and not $\hat{q}'$. Thus, when transforming from the reference frame of O to $\bar{O}$, the state vectors change, but the operators remain the same [Fonda and Ghirardi, 1970, p. 21].

3.3 Gauge notions

Finally, I turn to gauge theory. In this section, I will not explicitly discuss how the notion of different "gauge observers" is handled within this theory. This will be reserved for chapter 5. Instead, I will introduce the basic geometrical notions that are necessary before such a discussion can be held.

A distinctive feature of spacetime theories is that they can be formulated in terms of geometrical objects on a manifold. In order to facilitate the comparison of spacetime theories with gauge theories, it is advantageous to have a similar way of formulating the latter. However, because the "gauge degrees of freedom" are in so-called *internal spaces*, the mathematical setting of typical spacetime theories is not sufficient and more sophisticated tools are needed. The language of fibre bundles allows the internal degrees of freedom to be geometrized in a similar way to the external ones of spacetime theories, thereby facilitating the comparison.

The basic idea of a fibre bundle can be motivated in the following way. Consider some object with internal degrees of freedom, which means that some property ϕ of the object takes, at each point of spacetime, a value in some separate geometrical space I.[6] Thus, specifying the value of this property requires a map $\phi : \mathcal{M} \to I$, with $\phi(x)$ giving the value of the property at the point x of $\mathcal{M}$ [Isham, 1999, p. 199]. This is illustrated in figure 3.2.

However, a major (methodological) lesson of general relativity is that the objects described by a theory should not have properties that are compara-

[6] A spinor field would be an example of such an object. Later we will see that the phase of the electron can also be handled in this way. In order to incorporate electromagnetism, we will see that the phase of the wavefunction is actually determined via another geometrical space that encodes the interaction of the electron with the electromagnetic field.

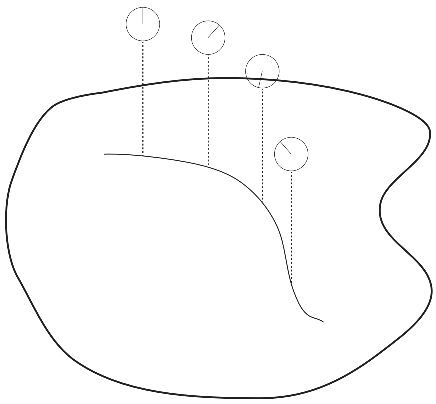

Figure 3.2: An example of a property ϕ that takes values in a clock-like space at each point along the curve.

ble "at a distance". In his early (unsuccessful) unification of gravitation and electromagnetism, Herman Weyl took this idea very seriously, and proposed that, as in GR the directions of vectors at distantly separated spacetime points cannot be compared as if they were at the same point, so the lengths of vectors at distantly separated spacetime points should also not be comparable [O'Rafeartaigh, 1997]. In the context of the present discussion, this means that the property ϕ takes values in different spaces for each spacetime point, thus $\phi(x) : \mathcal{M} \to I_x$. This entails that, if $x \neq x'$, then $\phi(x) \in I_x$, $\phi(x') \in I_{x'}$ and, in order to compare $\phi(x)$ and $\phi(x')$ some "comparing function" from I_x into $I_{x'}$ is necessary. This comparing function is called a *connection*.

In fact, these ideas have a very direct application in the more familiar context of spatial geometry. Consider the surface of a sphere, which is a two dimensional manifold $\mathcal{S}$. The tangent vectors to points in S each take values in separate tangent spaces $T_x\mathcal{S}$. Each $T_x\mathcal{S}$ is a two dimensional vector space. However, if I take a vector $\vec{v} \in T_x\mathcal{S}$ and another vector $\vec{v}' \in T_{x'}\mathcal{S}$ then I cannot compare the directions that $\vec{v}$ and $\vec{v'}$ point in. In other words, I cannot say if these two vectors are parallel or not. In order to do this, I need to define the notion of *parallel transport*, which means I need to specify some function $f : T_x\mathcal{S} \to T_{x'}\mathcal{S}$ which defines the notion of parallel. Thus, $\vec{v}$ is parallel to $\vec{v}'$ if $f(\vec{v}) = \vec{v}'$. In fact, f should be an isomorphism between $T_x\mathcal{S}$ and $T_{x'}\mathcal{S}$, and it is then called an *affine connection* [Dieks, 2011, p 117]. In general, one can define an infinitesimal connection, which specifies how a vector is parallel transported from one point to all the other points infinitesimally close to it. Comparing two distantly separated vectors then requires transporting the vector from the starting point to the end point in infinitesimal steps. Curiously, this entails that "parallelism" becomes a path dependent relation.

The geometry of the sphere can also be used to introduce the notion of a *bundle*. Consider "gluing" the space S and all the tangent spaces $T_x\mathcal{S}$ together into one "big" space, $T\mathcal{S}$. $T\mathcal{S}$ is called a *tangent bundle*. A vector field over $\mathcal{S}$ can be defined by picking out one vector from each of the tangent spaces $T_x\mathcal{S}$.[7] The result of such a process is called a *cross-section* of the tangent bundle, another notion that will be useful to us in the context of gauge theories [Isham, 1999, p. 201].

The tangent spaces considered in the case of the sphere were just one particular example of a possible space than can be "attached" to each point of a base manifold. Furthermore, tangent spaces in this sense are not usually referred to as internal spaces. For the internal spaces of gauge theory, we will want to attach other kinds of spaces, denoted F_x, to the base spacetime manifold $\mathcal{M}$. The spaces that we attach are called *fibres*, and we will require that they depend in some smooth differentiable way on position in the spacetime manifold. Formally, this translates into the requirement that the collection E of all the fibres F_x, $E = \bigcup_{x\in\mathcal{M}} F_x$ is itself a topological space (ibid, p.202). We can now give the general definition of a bundle.

Definition 8. [Isham, 1999, p. 202] A *bundle* is a triple $(E, \pi, \mathcal{M})$, where E and $\mathcal{M}$ are topological spaces, and $\pi : E \to \mathcal{M}$ is a continuous map.

E is called the *bundle* or *total* space, $\mathcal{M}$ is the *base* space and π is the *projection*. The fibres (the spaces that are attached to each point in the base space), are given by the inverse image $\pi^{-1}(\{x\}) = \{y|\pi(y) = x\}$. Note that π^{-1} is not a function, because it sends a point to a set of points. The projection is the crucial tool in relating the fibres to their associated base space points.

The bundles with which we will be concerned in this thesis will all share the property that their fibres are homeomorphic (in fact diffeomorphic because the fibres will be manifolds) to a common space F. F is then called the *fibre* of the bundle, and the bundle itself becomes a *fibre bundle*. An impression of a fibre bundle is given in figure 3.3 A fibre bundle is called trivial if it is isomorphic to the product bundle $(\mathcal{M} \times F, \mathrm{pr}_1, \mathcal{M})$ where for a point $(x, f) \in \mathcal{M} \times F$, $\mathrm{pr}_1(x, f) = x$ [Isham, 1999, p. 204]. Intuitively, a bundle is trivial when there are no "twists" in the way the fibres are attached to the base space. For instance, gluing the interval $[-1, 1]$ to each point of a circle S^1 gives the trivial bundle $(S^1 \times [-1, 1], \mathrm{pr}_1, S^1)$, depicted in figure 3.4.

A famous example of a non-trivial bundle is the Möbius strip. At first sight,

[7] However, there is no way to do this smoothly over the whole sphere. This is the problem of combing the hairy sphere.

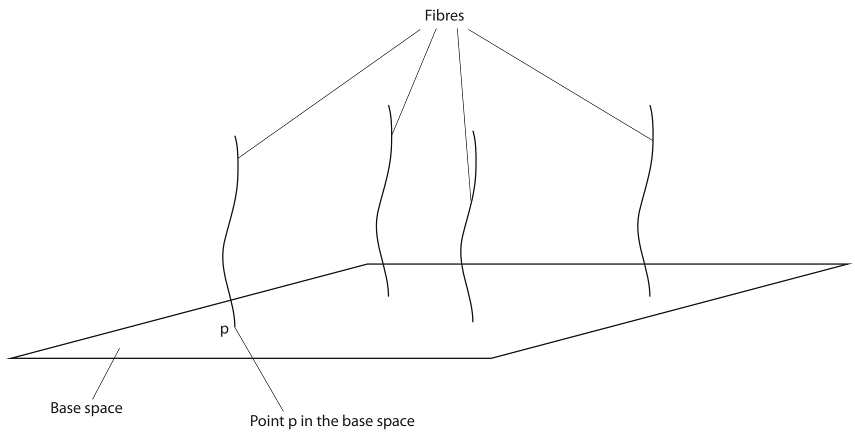

Figure 3.3: An impression of a fibre bundle (all the fibres are copies of each other), with only a few fibres shown explicitly.

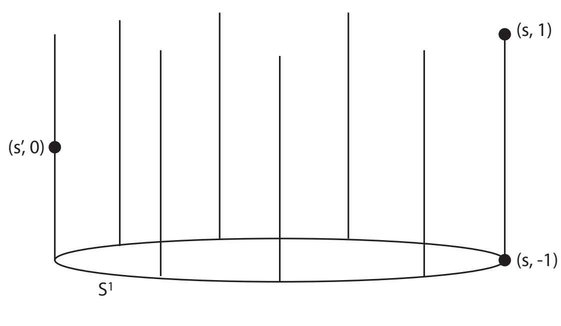

Figure 3.4: A representation of the trivial product bundle $(S^1 \times [-1, 1], \mathrm{pr}_1, S^1)$.

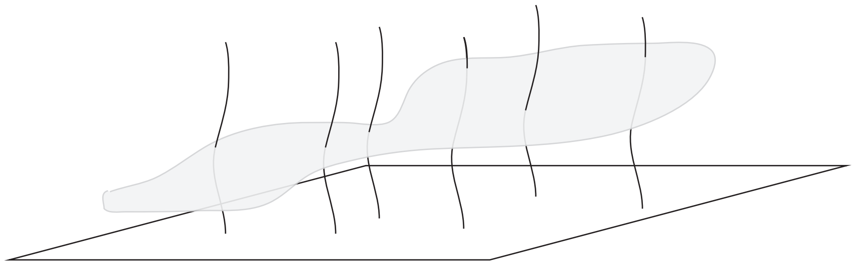

Figure 3.5: An impression of the cross-section of a fibre bundle. The grey surface cuts each fibre once and thus picks out a unique element of each fibre.

the Möbius strip might seem like it can be represented by the product bundle $(S^1 \times [-1, 1], \mathrm{pr}_1, S^1)$. The base space of the Möbius strip is indeed a circle, but the fibres must be attached in a more complicated way in order capture the way that it twists.

As we did earlier, it is possible to define the *cross-section* of a fibre bundle, by picking out one element of the fibre for each point in the manifold. This is illustrated in figure 3.5.[8]

Definition 9. [Isham, 1999, p. 207] A *cross-section* of the bundle $(E, \pi, \mathcal{M})$ is a map $\sigma : \mathcal{M} \to E$ such that the image $\sigma(x)$ of each point $x \in \mathcal{M}$ lies in the fibre $\pi^{-1}(\{x\})$ above x. More precisely, $\pi \circ \sigma = id_{\mathcal{M}}$, where $id_{\mathcal{M}}$ is the identity on $\mathcal{M}$, $id_{\mathcal{M}}(x) = x$.

In gauge theory, the main types of fibre bundles utilized are *principal fibre bundles*. A principal fibre bundle (also called principal bundle) is a fibre bundle with a Lie group G as fibre. Lie groups are differentiable manifolds, so the idea of attaching a Lie group to each point of spacetime falls neatly in line with the ideas presented above. E is called a *right-G space*, if an element $g \in G$ acts on a point $p \in E$ to give another point $p' \in E$ as $pg = p'$. The orbit of G at a point $p \in E$ is the set of points in E that can be reached by acting on p with G. In the case of a principal fibre bundle, the fibre above $x = \pi(p)$ is thus the orbit of G at p. If p is in the fibre above $x = \pi(p)$, and p'' is in the fibre above $x'' = \pi(p'')$, and $x \neq x''$, then there is no $g \in G$ such that $p'' = pg$. E/G is called the orbit space of the G-action on E [Isham, 1999, p. 221]. It is possible to define a function ρ such that if two points $p_1, p_2 \in E$ are in the orbit of G at p, then $\rho(p_1) = \rho(p_2) = \rho(p)$ is a point of E/G. Using these notions, the formal definition of a principal bundle can

[8] I will argue in chapter 5 that the cross-sections of certain fibre bundles, namely principal fibre bundles, are analogous to coordinate systems, in that they introduce a certain "point of view" of the bundle.

be given.

Definition 10. [Isham, 1999, p.221] A bundle $(E, \pi, \mathcal{M})$ is a principal bundle if E is a right G-space and it is isomorphic to $(E, \rho, E/G)$ where ρ is the usual projection map, and G acts freely on E, which means that each orbit is homeomorphic to G. This entails that G is the fibre of the bundle. G is called the *structure group* of the bundle.

I will now state a theorem that will be essential to all of our future uses of fibre bundles in gauge theory.

Theorem 2. *[Isham, 1999, p. 230] A principal G-bundle $(E, \pi, \mathcal{M})$ is trivial (i.e. isomorphic to $(\mathcal{M} \times G, \mathrm{pr}_1, \mathcal{M})$) if and only if it possesses a continuous cross-section.*

A better feeling for this theorem can be acquired by thinking back to the Möbius strip. In that case, it is not possible to define a continuous cross-section because a function from S^1 to $[-1, 1]$ does not return to its starting point after one period of the base space. Instead, it only returns to its starting point after two periods. For this reason, the cross-section of the Möbius strip is called anti-periodic.

Before presenting the final few definitions that we will need to use fibre bundles in our analysis of gauge theory, let me take stock of what we have achieved thus far. We started by imagining that there exists some physical object which has a property ϕ that takes values in some geometric space F. We accepted that the values of this property at distantly separated points should not be comparable (without defining some additional structure to enable a comparison). This forced us to stick a copy of the space F (called the fibre of the bundle) to each point x of spacetime. The copy of F at a point x we called the fibre above x. This process of attaching copies of F to points in the spacetime manifold $\mathcal{M}$ led us to define the notion of a fibre bundle, $(E, \pi, \mathcal{M})$. When the fibre is a Lie group G, then we have a principal fibre bundle.

Just as in the case of the vectors in the distantly separated vector spaces tangent to the surface of a sphere, we would like to have a way of comparing the property ϕ of the object when it changes its position in the base manifold (when it moves around in spacetime). In other words, we need a way of comparing points in the fibres above two distantly separated points x and x' of the base manifold $\mathcal{M}$. As we did in the case of the sphere, we will define a *connection* that allows such a comparison to be achieved by specifying the "parallel transport" of the object in infinitesimal steps from x to x'. A connection in a fibre bundle is a mathematical object that points from a point p in the fibre above x to points in the fibres above all the neighbouring points of x, thereby specifying which points in these neighbouring fibres are "the same as" (or "parallel to") p. Bernstein and Phillips suggest picturing a connection as a little "slope" attached to each point of a fibre that shows you how to go across to the next fibre (Bernstein and Phillips [1981], p.18). Following the slope corresponds to moving the object in spacetime while preserving the value of the property ϕ, and not following the slope corresponds to the property ϕ changing as the object moves in spacetime.

Providing a mathematically rigorous definition of a connection would take us too far afield, and therefore we will settle for trying to motivate the concept loosely. Consider the vector space T_pE at a point p of the total space E. Vectors in this tangent space will point from p to neighbouring points. All the vectors which point from p to other points in the same fibre as p form a subspace V_pE of T_pE called the *vertical subspace*. A connection is defined as follows.

Definition 11. [Isham, 1999, p. 254] A connection in a principal bundle $(E, \pi, \mathcal{M})$ with structure group G is a smooth assignment to each point $p \in E$ of a *horizontal subspace* H_pE of T_pE such that:

1. Any vector in T_pE can be decomposed uniquely into a sum of horizontal and vertical components.

2. H_pE is constant along the fibre to which p belongs (it is compatible with the right action of G on p).

In other words, a connection specifies which vectors point from a point in one fibre to a point in a neighbouring fibre without also moving up and down in that fibre. As a mathematical object, a connection can be represented as a Lie-algebra valued one-form ω. ω_p acts on vectors in T_pE. Given a vector $\tau \in T_pE$, if $\omega_p(\tau) = 0$, then $\tau \in H_pT$ which means that τ points from p to "the same" point in the neighbouring fibre of p.

The connection can be used to define the important notion of a *horizontal lift*. Roughly, a horizontal lift is a curve that moves through points in the fibres without moving up and down in the fibres. In other words, if τ is the tangent to the horizontal lift at a point p, then $\omega_p\tau = 0$.

Definition 12. [Isham, 1999, p. 263] Let α be a smooth curve that maps a closed interval $[a, b] \subset \mathbb{R}$ into $\mathcal{M}$. A *horizontal lift* of α is a curve $\alpha^\uparrow : [a, b] \to E$ which is horizontal ($\omega([\alpha^\uparrow]) = 0$) and such that $\pi(\alpha^\uparrow(t)) = \alpha(t)$ for all $t \in [a, b]$.

It is possible to use the horizontal lift to define parallel transport. Given an arbitrary curve $\alpha : [a, b] \to \mathcal{M}$ in the base space $\mathcal{M}$, the parallel transport of a point p in the fibre above $\alpha(a)$ to the fibre above $\alpha(b)$ is given by the point $p' \in \pi^{-1}(\alpha(b))$ reached by the horizontal lift of α which also passes

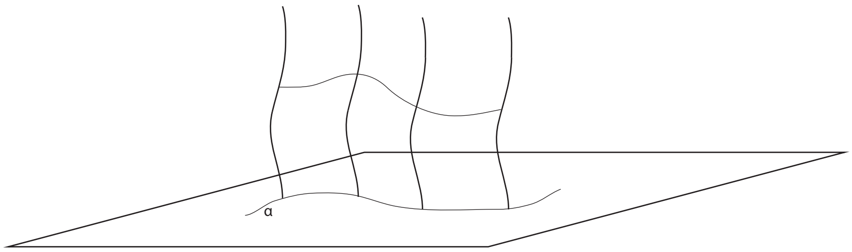

Figure 3.6: The arbitrary lift of a curve α in the base space into the total space. The connection will determine whether this curve is a horizontal lift or not.

through p (this is uniquely defined). The formal definition follows.

Definition 13. [Isham, 1999, p. 267] Let $\alpha : [a, b] \rightarrow \mathcal{M}$ be a curve in $\mathcal{M}$. The *parallel translation* along α is the map $\tau : \pi^{-1}(\{\alpha(a)\}) \rightarrow \pi^{-1}(\{\alpha(b)\})$ (from the fibre above $\alpha(a)$ to the fibre above $\alpha(b)$) obtained by associating with each point p in the fibre above $\alpha(a)$ the point $\alpha^{\uparrow}(b)$ in the fibre above $\alpha(b)$ where $\alpha^{\uparrow}$ is the unique horizontal lift of $\alpha(t)$ that passes through p at $t = a$.

The final concept that we need to define is that of *curvature*. In the case of the sphere, curvature is evident in the fact that, when a vector is parallel transported round a closed curve (a loop), it does not return to its original direction. Curvature at a point is defined by measuring the angular deviation between a vector and its parallel transport around an infinitesimal loop. In the case of fibre bundles, the notion of curvature can be defined in a similar way. Consider the horizontal lift $\alpha^{\uparrow}$ of a closed curve $\alpha : [a, b] \rightarrow \mathcal{M}$ (thus $\alpha(a) = \alpha(b)$). This can be used to define a map of the fibre above $\alpha(a)$ onto itself given by $p \rightarrow \tau(p)$, where τ is the parallel translation along α (see previous definition). In this case, because τ maps the fibre onto itself, an element of the structure group G can be associated with each loop. The element of G associated with an arbitrary loop α starting and ending at $\alpha(0)$ is called the *holonomy* of α. If you consider all possible loops in the base space $\mathcal{M}$ starting and ending at $\alpha(0)$ you obtain a map from the loop space of $\mathcal{M}$ into G, giving a subgroup of G called the *holonomy group* of the bundle at $\alpha(0)$ [Isham, 1999, p. 267]. The non-closure of the horizontal lift of a curve in the base space is illustrated in figure 3.7.

The curvature at a point of the bundle can therefore be defined in terms of the holonomies of infinitesimal loops starting and ending at the point. Mathematically, the curvature is a Lie-algebra valued two-form, which means that it maps a pair of vectors to an element of the Lie algebra. A final valuable feature of the holonomies of loops in $\mathcal{M}$ is that they provide information on the global topological properties of $\mathcal{M}$ (for instance, whether $\mathcal{M}$ is simply connected or not, etc...). This will be important in our future discussions of the Aharonov-Bohm experiment.

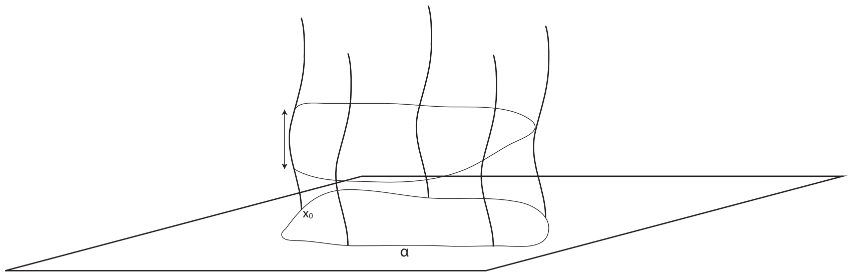

Figure 3.7: A non trivial holonomy. The horizontal lift of the closed curve α in the base space does not close (failure of closure is indicated above the point x_0).

In chapter 5, these mathematical tools will be put to use to understand how transformations are defined and interpreted in the theory of a non-relativistic quantum particle interacting with an electromagnetic field.

Chapter 4

The role of passive transformations

In this chapter, I hope to provide an answer to the following question posed by Paul Teller [Teller, 2000, p. 479]:

> In what way, or to what extent, does the present approach to electromagnetism [the electromagnetic potential as a connection defining parallel transport] present electromagnetism itself as a geometrical phenomenon in something like the way that gravitation is presented as a geometrical phenomenon in general relativity?

In the next chapter, I will show how the electromagnetic potential can be represented in terms of a connection on a principal fibre bundle. However, before fully exploiting the fibre bundle machinery, I will give a full presentation of the "symmetry dictates interaction" role of gauge symmetry, also known as the "gauge principle". I will compare this inference from symmetry to interaction with the role of coordinate transformations and the equivalence principle in GR. I will define an analogue of the inertial coordinate system in gauge theory, called the inertial gauge. I will show that electromagnetism determines the local inertial gauge in the same way that gravitation determines the local inertial coordinate system.

4.1 Introduction: The geometrical programme

With the advent of general relativity, Einstein is often credited with having founded what Cao refers to as the "geometrical programme" [Cao, 1990, p. 117]. This programme prescribes a certain methodology for physical research into the fundamental forces of nature. In the broadest possible terms, it aims to describe the forces of nature in terms of the geometrical properties, like curvature or topology, of some geometrical space, which are encoded by geometrical objects defined on the space, like metrics and affine connections. In the case of GR, a field equation relates the geometrical objects, defined on spacetime, to the sources of the "force", which are mass-stress-energy.[1] As a result, the geometry of spacetime becomes a *dynamical* part of the physics, rather than a backdrop against which events unfold.

After the success of general relativity, it was natural to ask whether a similar geometrization of the other force known at the time, electromagnetism, was possible. Attempts in this direction led to Weyl's ill-fated first gauge theory, Kaluza-Klein theories, and occupied Einstein for the rest of his life with work on so-called unified field theories. Although most of these early trials were unsuccessful, they sowed the seeds from which the undeniably successful gauge theories would grow. Today, these theories describe three of the four fundamental forces: the strong force, the weak force and the electromagnetic force [Weinberg, 1977, p. 32]. However, the status of gravity as a gauge theory in the same sense as that of the other three forces is controversial, and the debate is ongoing. Part of the problem is the extent to which the geometrization of the other three forces can be said to be achieved "in the same way" as that of gravitation.

It is a widely held view that the universality of gravitation is crucial to the possibility of giving it a geometrical formulation. "Universality" here refers to the fact that all particles follow geodesics of spacetime regardless of their constitution (mass, charge etc...). In this vein, Brown and Pooley stress that "mass is not a coupling constant" because "it does not indicate the strength of the particle's coupling to the connection" [Brown and Pooley, 2006, p. 72]. As is well known, if released in a vacuum in gravitational field, a large rock and a small feather will fall with the same acceleration. In contrast, the charge of a particle *is* a coupling constant to the electromagnetic field. Two particles of equal mass but of different charge will behave differently in identical electromagnetic and gravitational fields.

In the face of this disanalogy between mass and charge, any claim to the "geometrization" electromagnetism would appear thin. The issue obviously turns on the view of "geometry" that one adopts. Traditionally and roughly speaking, geometry is the mathematical study of the properties of physical space familiar from everyday experience. With the developments of Minkowski, geometry was extended to include time, and thus arose the concept of spacetime. All physical bodies live and move in spacetime, and because all these bodies are affected in the same way by gravitation, it is

[1] I place the word "force" in scare quotes because, once the geometrical programme has been adopted, talk of forces in the Newtonian sense becomes obsolete.

possible to encode the effects of gravitation in the very structure of spacetime itself.[2] However, the argument goes, not all bodies are affected in the same way by electromagnetism, and therefore electromagnetism cannot be built into the structure of spacetime. This suggests that electromagnetism cannot be geometrized if one limits oneself to spacetime geometry. However, if the notion of geometry can be extended to include yet more features, it is conceivable that electromagnetism could be represented geometrically with the help of these additions. This is precisely how the geometrization of electromagnetism is achieved in the fibre bundle formalism.

Such reasoning raises two questions: 1) If a so-called geometric formulation of electromagnetism is not based on the geometry of spacetime, to what extent can it really be considered as geometrized "in the same way" as gravitation? 2) When does a piece of mathematics deserve to be called geometry or not? And is this important? I will answer these questions in turn. When it is claimed that electromagnetism is geometrized "in the same way" as gravity, this must not be interpreted as a claim that electromagnetism can also be incorporated into the structure of spacetime itself. As Trautman points out, the unification sought by advocates of "connections on fibre bundles" is "considerably different from Einstein's own attempts", precisely because Einstein limited his geometrical tools to modifications of Riemannian geometry [Trautman, 1980, p. 288]. Kaluza-Klein theories were similarly conservative, seeking their unifications in higher-dimensional *spacetime* rather than in more exotic spaces. In this respect, the geometry of fibre bundles is clearly different from that of GR. On the other hand, as was shown in the previous chapter, the same geometrical concepts of connections, curvature, parallel transport, covariant derivative etc... are employed in the mathematics of fibre bundles. This answers the questions posed above, and justifies calling fibre bundles geometrical structures. It is because these same concepts can be used to represent electromagnetism (albeit in new geometrical spaces rather spacetime) that it can be claimed that electromagnetism (and the other forces) can be geometrized "in the same way" as gravity. Furthermore, the mathematics of fibre bundles does lend itself to a pictorial representation (as we introduced in chapter 3 and will develop further in chapter 5) with a distinct geometrical flavour.

Unfortunately, just as some worries about the similarities between the geometrical formulations of GR and the other gauge theories are dispelled, others appear. As one looks closer at the precise mathematical structures that are used to represent the various forces, it becomes clear that there are substantial technical differences which prevent one from comfortably being able to claim that GR is a gauge theory like the others. One way to sum

[2] I gloss over any phenomena that might not fit neatly into a spacetime picture, such as the consequences of quantum entanglement for instance.

up these technical differences is to point to the fact that the gauge theories of the electromagnetic, strong and weak forces are all Yang-Mills theories, whereas GR is not. As Earman explains, a defining feature of Yang-Mills theories is the closure of the Lie algebra of the constraints in the constrained Hamiltonian formalism. In this formalism, the Lie algebra of the constraints of GR is not closed [Earman, 2002, p. 217].

Alternatively, one might try to give an explicit fibre bundle formulation of GR. For Trautman, a gauge theory is "any physical theory of a dynamical variable which, at the classical level, may be identified with a connection on a principal bundle" [Trautman, 1980, p. 306]. Thus, if it were possible to reformulate GR in this way, its status as a gauge theory like the others would be further supported. However, this strategy also runs into technical difficulties, and Trautman admits that even when given a fibre bundle formulation, "gravitation is different from other gauge theories". As the source of these differences, he points to the "soldering of the bundle of linear frames LM to the base manifold M" [Trautman, 1980, p. 299]. "Soldering" is a technical term of gauge theory which can be explicated as follows. As explained in chapter 3, a fibre bundle consists of a base space M with some other space called a fibre attached to each point of M. In the theories of the electromagnetic, strong and weak forces, the fibre is an internal space that is unrelated to spacetime. However, in the fibre bundle formulation of GR, the geometry of the fibres encodes the geometry of the base space itself [Lyre, 2000, p 6]. This "intimate relation" between the base space and the fibres in the case of GR differentiates it from the other gauge theories.

Apart from the technicalities, there are also more conceptual disagreements about the gauge nature of gravity. These often focus on the identification of the "gauge group" of GR. The gauge theories of the strong, weak and electromagnetic forces are explicitly constructed around the groups SU(3), SU(2) and U(1), corresponding respectively to the three forces. However, in the case of gravity the situation is somewhat confused. As Redhead notes [Redhead, 2003, p. 134]

> So, there is considerable confusion 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!

General coordinate transformations, so called "passive transformations", are often considered to have an active equivalent, namely diffeomorphisms. Motivated by the hole argument and the general covariance of GR, Earman and Norton urge that general coordinate transformations (and their associated diffeomorphisms) represent the "gauge freedom" of GR. Norton urges us to recognize the "central importance of general covariance as a gauge freedom of general relativity" [Norton, 2003, p. 110]. Earman even goes so far as to dismiss the fibre bundle approach as "glitzy", and advocates the constrained Hamiltonian formalism as the "non-question begging and systematic way to identify gauge freedom" [Earman, 2002, p. 212]. As if the waters were not murky enough, Weinstein argues that it is "clearly misguided" to "think of the diffeomorphism group as a gauge group" [Weinstein, 1999, p. 8].

I believe that much of the confusion is down to an equivocation over the notions of "gauge group" and "gauge freedom". Very different accounts of gauge develop depending on where the emphasis is placed. The notion of a "gauge group" is generally used in the context of a unified methodology for geometrizing fundamental forces. On the other hand, "gauge freedom" is employed in the context of discussions about redundancies in physical theories. These two facets feed off each other, but the debates in which they feature ultimately have different motivations. The methodological aspect of gauge is primarily driven by a desire for unification and conceptual harmony between the different fundamental physical theories. On the other hand, the debate about redundancy is focused on the notions of "physicality" and "reality", and is connected to Redhead's concept of surplus structure. In this context, the main concern is with identifying the parts of physical theories that have correlates in the world, and it is believed that redundant elements of a theory are not well suited for this role. In this thesis, my focus is on the methodological and unifying aspects of gauge. Moreover, by highlighting the importance of gauge in those contexts, I wish to discourage the tendency to interpret the gauge elements of a theory as redundancies that can best be done without.

In the first part of this chapter, my objective will be to support the claim that gravitation and electromagnetism can be geometrized in the same way by comparing the two principles, the equivalence principle and the gauge principle, around which the geometrization is based. I will give an interpretation of the equivalence principle which makes its relation to the gauge principle evident. I will note that the objections raised against the logical roles of the two principles in their respective theories are very similar. My conclusion will be that both principles are responsible for providing the *possibility* of giving a geometrical account of the forces of nature.

In the second part, I will present a consequence of my interpretation of the equivalence principle for the thorny issue of the equivalence of coordinate systems in GR, and clarify the role of the Einstein field equations (EFE's) within this account. The main lines of reasoning are inspired by a proposal from Dieks for reinterpreting the sense in which GR can be said to extend the relativity principle of special relativity [Dieks, 2006]. By appealing to the Aharonov-Bohm effect, I will then argue that a very similar story can be told in quantum-electromagnetism (Q-EM). I show how the equivalence principle has a precise analogue in Q-EM, and make a suggestion for the form of the field equations. We will see that these field equations are in fact the holonomies of the principal fibre bundle that represents the electromagnetic field. I will note an important geometrical difference between GR and Q-EM: in GR, it is the curvature of spacetime that encodes the presence of a gravitational source whereas in Q-EM it is the topology that encodes the presence of a phase-shifting source.

4.2 Part I: From symmetry principles to geometry

The equivalence principle is a notoriously slippery ingredient of GR. It has many different formulations, and opinions on its importance within the theory cover the entire range from fundamental to dispensable. For instance, in Weinberg's textbook *Gravitation and Cosmology*, the principle occupies pride of place. However, others are less accommodating. Synge claims that he never understood the principle, (Norton [1985], p.243) and is scathing about its role within the final theory (Synge [1960], p.xi):

> The Principle of Equivalence performed the essential office of midwife at the birth of general relativity, but, as Einstein remarked, the infant would never have got beyond its long-clothes had it not been for Minkowski's concept. I suggest that the midwife be now buried with appropriate honours and the facts of absolute spacetime faced.

In what follows, I will present the equivalence principle as a way of interpreting the mathematical formalism that is indispensable for incorporating gravity into the geometrical structure of spacetime. In fact, it seems that Einstein himself may have had a similar understanding of the principle.

4.2.1 Formulating the equivalence principle

Most people's first encounter with the equivalence principle is probably in a science class demonstration that, in a vacuum tube, a feather and a rock fall at the same rate. In this context, the principle is taken to state the equality of inertial and gravitational mass. In a general relativity class, the principle is often given in its infinitesimal form (Weinberg [1972], p.68)

> At every spacetime point in an arbitrary gravitational field it is possible to choose a "locally inertial coordinate system" such that, within a sufficiently small region of the point in question, the laws of nature take the same form as in unaccelerated Cartesian coordinate systems in the absence of gravitation.

In popular science books, the principle is often associated with Einstein's lift thought experiment (as presented in chapter 1). In agreement with Einstein's 1916 introduction to the theory, a crucial feature of all formulations of the equivalence principle is that they point to the indistinguishability of the physical consequences of being at rest in a gravitational field, or in a state of "really" accelerated motion (Einstein [1997], p.150). Precisely because of the impossibility of (locally) distinguishing such physical effects, the one can be used to cancel the other. In other words, an observer free-falling in a gravitational field can be considered to be "at rest". In this way, Norton believes that, for Einstein, "the basic assertion of the principle of equivalence is that "one may treat K' [the accelerated observer] as at rest" (Norton [1985], p.206).

The possibility of treating accelerated observers as at rest can also be understood mathematically. Representing the laws of physics in non-inertial (accelerated) coordinate systems requires the appearance of new terms in the equations of motion. These new terms are often referred to as "inertial" or "fictitious" forces. Thus, the inertial equations of motion

$$\frac{\mathrm{d}^2 x^\lambda}{\mathrm{d}t^2} = 0 \tag{4.1}$$

become

$$\frac{\mathrm{d}^2 x^\lambda}{\mathrm{d}t^2} + \Gamma^\lambda_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}t}\frac{\mathrm{d}x^\nu}{\mathrm{d}t} = 0 \tag{4.2}$$

in a non-inertial coordinate system. The $\Gamma^\lambda_{\mu\nu}$ are the Christoffel symbols, which, in the case of a rotating coordinate system for example, represent the centripetal and Coriolis forces. In this mathematical setting, Einstein's suggestion becomes to interpret the non-vanishing Christoffel symbols in a particular coordinate system as a gravitational field, and thereby to restore the possibility of considering that coordinate system as inertial. As Janssen puts it (Janssen [2011], p.2)

> Two observers in non-uniform motion with respect to one another can both claim to be at rest as long as they agree to disagree about whether or not there is a gravitational field.

Moreover, Einstein was able to show that, in Minkowski spacetime, the equations of motion for a uniformly accelerated observer could always be written in a form that closely resembled the equation for a particle moving under the influence of a Newtonian gravitational field (Norton [1985], p.217).[3] This further convinced him that the Christoffel symbols appearing in the equations of motion for non-inertial coordinate systems could be seen as closely related to gravitational effects. In his 1916 introduction to the general theory of relativity, he goes so far as to claim (Einstein [1997], p.150)

> It will be seen from these reflexions that in pursuing the general theory of relativity we shall be led to a theory of gravitation, since we are able to "produce" a gravitational field merely by changing the system of coordinates.

The problem with this approach is that it runs the risk of confusing coordinate effects with physical effects. If we believe that gravitational fields must have sources, then it is very unsatisfactory to have a theory which allows gravitational fields to be produced by changing coordinates, because a change of coordinates clearly cannot bring a source into existence. This is the essence of Synge's rejection of the equivalence principle (Norton [1985], p.243)

> Does it [the equivalence principle] mean that the effects of a gravitational field are indistinguishable from the effects of an observer's acceleration? If so, it is false. In Einstein's theory, either there is a gravitational field or there is none according as the Riemann tensor does or does not vanish.

The view that it is the curvature of spacetime as encoded by the non-vanishing of the Riemann curvature tensor that signals the presence or absence of gravitational effects is the currently accepted interpretation of general relativity. Precisely because of its tensorial properties, the Riemann curvature tensor cannot be made to appear or vanish by a coordinate transformation, and therefore qualifies as a more "objective" characterization of the presence (or absence) of gravitation.

Apart from their coordinate dependence, there is another unconvincing feature of the Christoffel symbols in Minkowski spacetime that prevents their interpretation as representing gravitational effects. Consider once again the case of a uniformly accelerating observer. In this case, the Christoffel symbols would represent a uniform gravitational field, in the sense that all the particles in the vicinity of the observer would be accelerated in parallel (there would be no tidal forces). However, apart from the rather unphysical case of a infinite plate with a uniform mass density, gravitational fields are always non-uniform because they are caused by finitely extended sources. Christoffel symbols with such a property cannot be created purely from a coordinate transformation.

[3] More specifically, Einstein was able to derive an equation that looked like acceleration = - gradiant of scalar field

Nevertheless, a uniform gravitational field can perhaps favourably be interpreted as an approximation to a real gravitational field. For instance, one that would work well for computing the trajectories of projectiles traveling only short distances near the surface of the earth. As Norton points out, such effects, which can be produced by the equivalence principle, "yield a special case of the gravitational field, whose properties are then generalized in a natural way to arrive at a general theory of gravitation" (Norton [1985], p.229).

I believe this remark is crucial to understanding the role that the principle plays in the logical structure of general relativity. It emphasizes how the Christoffel symbols, as generated by coordinate transformations, can be interpreted as encoding effects very similar to those of gravitation. The natural extension of this realization is to ask what properties the Christoffel symbols would have to have to actually represent real gravitational fields with sources. The answer, that Einstein arrived at with much difficulty, is that they must represent a curved (non-euclidean) spacetime. Of course, Christoffel symbols with this property cannot be created by pure coordinate transformations. However, by adding extra equations to the theory which they must obey, they can acquire this property. In fact, the extra equations are the Einstein field equations, and they determine the metric, which in turn determines the Christoffel symbols by the equation

$$\Gamma^{\mu}_{\sigma\rho} = \frac{1}{2} g^{\nu\mu} \left(\frac{\partial g_{\rho\nu}}{\partial x^{\sigma}} + \frac{\partial g_{\sigma\nu}}{\partial x^{\rho}} - \frac{\partial g_{\rho\sigma}}{\partial x^{\nu}} \right) \tag{4.3}$$

In Einstein's final theory, the Christoffel symbols thus play a dual role: they encode both the "inertial forces" present in a non-inertial coordinate system and the effects of gravitation. This has led some, such as Janssen, to suggest that in GR there appears a new unified "inertio-gravitational" field, which splits differently into inertial and gravitational components depending on the coordinate system (Janssen [2011], p.3). In this respect, this field is similar to the electromagnetic field in special-relativity, whose split into electric and magnetic components depends on the choice of a particular inertial coordinate system.

I believe that the above considerations can even partially elucidate the mystery of the role of "surplus structure" (as formulated by Redhead). In Minkowski spacetime, the Christoffel symbols appear as a form of surplus structure, only serving to compensate for the choice of a non-inertial coordinate system. However, once their presence in this context is discovered, it becomes possible to ask what extra properties they could have, and whether such additional features could be exploited to represent physical features of the world. In a sense, the methodology here is first to introduce some "slack" into the theory, and then to "take it up" by exploiting it in new creative ways. We will see in the next section that the gauge principle works in a very similar manner.

4.2.2 Formulating the gauge principle

I will first present a standard formulation of the gauge principle in the context of non-relativistic quantum-electromagnetism and then a relativistic version for the Dirac field. Afterwards I will explain how it can be seen to work analogously to the equivalence principle as formulated in the previous section. The version of the non-relativistic gauge principle presented here follows closely that of Aitchison and Hey [Aitchison and Hey, 1989, p. 51].

Consider an electron "moving freely" (in the absence of a potential) described by the Schrödinger equation (I use natural units, $\hbar = c = 1$)

$$\frac{-1}{2m}\nabla^2\psi(\mathbf{x},t) = i\frac{\partial\psi(\mathbf{x},t)}{\partial t} \tag{4.4}$$

This equation is invariant under the "global" transformation

$$\psi(\mathbf{x},t) \rightarrow \psi'(\mathbf{x},t) = e^{iq\chi}\psi(\mathbf{x},t) \tag{4.5}$$

where q and χ are constants.[4] The transformation is called "global" because χ is not a function of space and time. By "invariant", we mean that if the new value of the wavefunction $\psi'(\mathbf{x},t)$ is plugged into equation (4.4), all the "extra pieces" will cancel out, and the equation returns to its standard form (4.4).

Now comes the "magic" of the gauge principle. Consider the transformation properties of the Schrödinger equation under the "local" transformation

$$\psi(\mathbf{x},t) \rightarrow \psi'(\mathbf{x},t) = e^{iq\chi(\mathbf{x},t)}\psi(\mathbf{x},t) \tag{4.6}$$

where $\chi(\mathbf{x},t)$ is now a function of space and time, and thus feels the action of both derivative operators ∇^2 and $\frac{\partial}{\partial t}$. The Schrödinger equation is no longer invariant under this transformation because extra pieces which do not cancel appear from the action of the derivative operators on $\chi(\mathbf{x},t)$. In order to make (4.4) invariant under the local transformation (4.6), we can introduce two new fields into the equation with the correct transformation properties. These fields are a vector field $\mathbf{A}$ and a scalar field ϕ with the transformation properties

[4]The significance of q will become apparent later.

$$\mathbf{A} \to \mathbf{A}' = \mathbf{A} + \nabla\chi(\mathbf{x}, t) \tag{4.7}$$

$$\phi \to \phi' = \phi - \frac{\partial\chi(\mathbf{x}, t)}{\partial t} \tag{4.8}$$

After the introduction of these two fields, the Schrödinger equation becomes

$$\frac{-1}{2m}(\nabla - iq\mathbf{A})^2\psi'(\mathbf{x}, t) + q\phi\psi'(\mathbf{x}, t) = i\frac{\partial\psi'(\mathbf{x}, t)}{\partial t} \tag{4.9}$$

which is in fact the equation for a particle of charge q interacting with an electromagnetic field, called a gauge field, represented by the vector potential $A^\mu = (\phi, \mathbf{A})$. Note that the transformation properties of the $\mathbf{A}$ and ϕ fields are precisely the gauge transformations of the electromagnetic vector potential. In this way, it seems that the form of the interacting theory can be derived by demanding that the Schrödinger equation be invariant under the local transformation (4.6). This is an example of what Wigner called a “dynamical symmetry”, and the process which Yang termed “symmetry dictates interaction” (see chapter 2). The interacting theories of the strong and weak forces are obtained in a similar way, by demanding the invariance of a particular Lagrangian under a particular local transformation.[5]

Because the Schrödinger equation is non-relativistic, the necessary changes that must be imposed on it to make it invariant under the local transformation (4.6) can seem somewhat cumbersome. However, when a theory can be given a Lagrangian formulation, the gauge principle can be applied in a smoother fashion. For instance, consider the Dirac Lagrangian for a spinor field Ψ of mass m

$$\mathcal{L}_{Dirac} = \bar{\Psi}(i\gamma^\mu\partial_\mu - m)\Psi \tag{4.10}$$

This equation is not invariant under the local transformation

$$\Psi(\mathbf{x}, t) \to \Psi'(\mathbf{x}, t) = e^{iq\chi(\mathbf{x},t)}\Psi(\mathbf{x}, t) \tag{4.11}$$

$$\bar{\Psi}(\mathbf{x}, t) \to \bar{\Psi}'(\mathbf{x}, t) = e^{-iq\chi(\mathbf{x},t)}\bar{\Psi}(\mathbf{x}, t) \tag{4.12}$$

In order to make it invariant, it is possible to replace the ordinary partial derivative with a covariant derivative

[5] In the case of the strong and weak force, the relevant transformations are elements of a non-Abelian group, which means they do not commute. This has consequences for the mathematical form of the symmetry restoring gauge field that must subsequently be introduced.

$$\partial_\mu \to D_\mu = \partial_\mu + iqA_\mu \tag{4.13}$$

Giving the new Lagrangian

$$\mathcal{L}_{Dirac_{int}} = \bar{\Psi}(i\gamma^\mu(\partial_\mu + iqA_\mu) - m)\Psi \tag{4.14}$$

which represents the theory of a massive charged Dirac field interacting with an electromagnetic field (in fact some terms are missing, as we will see shortly). It it also sometimes written as

$$\mathcal{L}_{Dirac_{int}} = \bar{\Psi}(i\gamma^\mu\partial_\mu - m)\Psi - qA_\mu\bar{\Psi}\gamma^\mu\Psi \tag{4.15}$$

to make the new interaction (coupling) term explicit. The form of this term is known as "minimal coupling". The role of the covariant derivative will be further discussed in the context of the fibre bundle formulation of non-relativistic quantum-electromagnetism.

Criticism of the gauge principle in the literature is abundant. A first problematic aspect is the inclusion of the charge q in the initial global transformation (4.5). It is this move which forces the further appearance of q in the interacting Schrödinger equation (4.9). The fact that such a charge is included "from the beginning", suggests that the gauge principle doesn't produce the interacting equation out of nothing, but instead seems to already know where it is going. Aitchison and Hey argue that by combining the absolute conservation of charge with the claim that only particles of equal charge can exhibit quantum mechanical interference effects one can justify the inclusion a coupling constant in the transformations (4.5) and (4.6) [Aitchison and Hey, 1989, p. 55]. However, they admit that it is difficult to justify why q should represent charge and not some other quantum number. The artificial inclusion of the coupling constant q is certainly problematic, but the main criticisms of the gauge principle usually focus on other issues.

In the literature, there are four main objections to interpreting the gauge principle as generating an interacting theory from a non-interacting one: (1) the demand for invariance under "local" transformations is not justified (2) the form of the interacting Lagrangian is not determined uniquely (3) the theory obtained from the gauge principle actually contains no interactions (4) the gauge field doesn't satisfy the action-reaction principle. I will address each of these concerns in turn.

Brown puts the first objection in the form of the question "why should we be interested in a gauge-covariant version of (4.10) in the first place?" and draws an analogy to the requirement of general covariance in GR [Brown, 1999, p. 7]. In the next chapter, I will further defend the analogy between gauge transformations and coordinate transformations, and therefore I believe that a similar analysis of the invariance requirement can (and should) be given in both cases. However, as is well known, it remains highly problematic what the correct analysis should be.

Perhaps a better approach to the question is to ask why the notion of a covariant derivative, in other words, a means of comparing the phase of the wavefunction at distantly separated points, should be introduced. This sense of locality is precisely what Weyl considered to be one of the most important features of general relativity, and therefore he was one of the first to suggest it should be extended to quantum theory. Still, this sense of locality remains difficult to justify. Ultimately, I believe that requiring an explicit rule for parallel transport is an essential technique in exploiting geometrical properties to represent features of the world. We will see this argument returning later in this section.

The second objection targets the inability of the requirement of gauge-invariance to uniquely specify the form of the interacting Lagrangian (4.15). As Martin explains, there are many other terms that could be added to the Dirac equation to make it invariant under gauge transformations. It is sometimes argued that the minimal coupling of (4.15) can be justified by claiming it is the "simplest, renormalizable, Lorentz and gauge invariant Lagrangian yielding second order equations of motion for the coupled system" [Martin, 2002, p. 228]. However, by appealing to the geometrical interpretation of the covariant derivative, I believe that a more appealing answer is at hand. Consider first the case of GR. When we demand a version of the equation of motion for a massive particle (equation (4.1)) that is valid in all coordinate systems, we unproblematically arrive at the standard geodesic equation (4.2). Of course, we could write down any number of other generally covariant expressions, but they would not be relevant to describing the free-fall of a massive particle. Geometrically, we are interested in an equation for the vanishing of the covariant derivative. As we will see later, exactly the same line of reasoning can be performed for a quantum particle. More precisely, we will describe the "phase-motion" of a quantum particle in terms of the vanishing of the covariant derivative of a specific mathematical quantity. The form of the interaction is then determined by the general way in which one defines covariant derivatives.

The third and fourth objections are the most interesting because they relate directly to our earlier discussion of the principle of equivalence. The third objection claims that the form of the A_μ field required by the gauge principle does not correspond to a physically different situation from that described by the non-interacting Schrödinger equation. From this, it infers that the gauge principle is not sufficient to generate a truly dynamic interacting theory of a quantum system with an electromagnetic field. In order to appreciate the force of this objection, we must show that a joint gauge transformation of the wavefunction and the electromagnetic potential as in (2.9) results in no physical change to the system. In the classical case, it was sufficient to show that a gauge transformation left the electric and magnetic fields unchanged. However, in the quantum case, we must also show that the transformation does not affect the state of the quantum particle. As we will see when we discuss the Aharonov-Bohm effect later in this chapter, transformations of the phase of the wavefunction will have empirical consequences, manifested as changes in the interference pattern, if they affect the integrability of the wavefunction. The wavefunction is said to be integrable if the phase shift it undergoes between two space(time) points is independent of the path traveled between the points. It was shown by Dirac that integrable phase factors correspond to harmless redefinitions of the momentum operator $\hat{p}$ [Aitchison and Hey, 1989, p. 57]. We will have more to say about the effects of phase shifts on the integrability of the wavefunction when we discuss active gauge transformations in chapter 5.

In order to show that a joint gauge transformation of the wavefunction and the electromagnetic potential has no physical consequences, we must show that it leaves the electric and magnetic fields unchanged and that it does not affect the integrability of the wavefunction. Given a force free quantum particle that obeys (4.4), the best that the gauge principle can do is require the introduction of a potential of the form $A^\mu = (\phi, \mathbf{A}) = (-\frac{\partial}{\partial t}\chi(\mathbf{x},t), \nabla\chi(\mathbf{x},t))$. First we must show that the electric and magnetic fields

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{4.16}$$

$$\mathbf{E} = -\nabla\phi - \frac{\partial}{\partial t}\mathbf{A} \tag{4.17}$$

are not affected by a vector potential of this kind. The results are

$$\mathbf{B} = \nabla \times \nabla\chi(\mathbf{x},t) = 0 \tag{4.18}$$

$$\mathbf{E} = -\nabla(-\frac{\partial}{\partial t}\chi(\mathbf{x},t)) - \frac{\partial}{\partial t}\nabla\chi(\mathbf{x},t) = 0 \tag{4.19}$$

Furthermore, solving the Schrödinger equation for a potential of this form gives [Aitchison and Hey, 1989, p. 57]

$$\psi = e^{iq\int_{-\infty}^{x}\nabla\chi(\mathbf{x},t)\cdot d\mathbf{l}}\psi(\chi = 0) \tag{4.20}$$

where d$\mathbf{l}$ is an infinitesimal vector tangent to the path from $-\infty$ to $\mathbf{x}$ along which the integral takes place and $\psi(\chi = 0)$ is the solution to the force-free Schrödinger equation. Because the integrand is of the form $\mathbf{A} = \nabla\chi(\mathbf{x}, t)$, the phase of the wavefunction at an arbitrary point x is indeed integrable. The interference pattern will therefore be left unchanged by a gauge transformation of the form (2.9). The physical situation described by (4.9) with $A^\mu = (-\frac{\partial}{\partial t}\chi(\mathbf{x}, t), \nabla\chi(\mathbf{x}, t))$ is therefore physically indistinguishable from the physical situation described by the force-free Schrödinger equation (4.4).

The above arguments show that demanding the invariance of the Schrödinger equation under local gauge transformations of the form (4.6) *does* require the introduction of a new mathematical structure into the theory, namely the A_μ field, but that this structure *does not* correspond to any new physical effects. Instead it simply seems to compensate for unusual choices of phase convention. The flavour of this objection is very similar to that of Synge's against Einstein's version of the equivalence principle. When taking non-inertial coordinate systems into consideration, Einstein was led to introduce the Christoffel symbols into the equations of motion for force-free particles. However, these Christoffel symbols simply encode the presence of fictitious forces resulting from the choice of a non-inertial coordinate system. As Synge pointed out, the presence of a real physical force, one with a source, is an observer independent matter, and therefore cannot be affected by the choice of a coordinate system. The same must hold in Q-EM. A change of coordinate system cannot bring a force into existence, and neither can a change of phase convention. In the case of GR, we argued that Einstein's great achievement was to realize that the Christoffel symbols could be made to represent gravitational effects if they obeyed additional requirements, ones not dictated purely by general covariance.

In the case of Q-EM, two conditions conspire to obscure the analogy of the gauge principle with the role of the equivalence principle in GR. The first is a contingent fact about the way in which both theories were discovered. The second is related to the nature of our experience as human beings. When Einstein was working toward the general theory of relativity, he did not know what the final theory would look like. Therefore, when he started to consider non-inertial coordinate systems, he did not immediately recognize the fact that the Christoffel symbols could also be used to encode proper gravitational effects (in Synge's sense). It was only after much work that he realized that the Christoffel symbols could describe such effects if they gave rise to a spacetime geometry with non-zero curvature. In the case of Q-EM however, the situation is precisely reversed. It was already known what the final theory looked like before the case of "general" phase conventions, instantiated by the transformations (4.6), were considered. Once it was realized that the introduction of a field that looked (formally) exactly like A_μ restored invariance under local gauge transformations, there was no problem in seeing this field as a special case of the electromagnetic potential. The difference can be summed up as follows. Einstein saw the ability to represent gravitational effects as a special extension of the already known role of the Christoffel symbols in accounting for non-inertial coordinate systems. On the other hand, physicists saw the "impotent" $A^\mu = (-\frac{\partial}{\partial t}\chi(\mathbf{x}, t), \nabla\chi(\mathbf{x}, t))$ field introduced by the gauge principle as a certain restricted version of the already known electromagnetic potential.

The second condition is the very different nature of the fictitious forces that appear in non-inertial coordinate systems, and those that appear in "non-inertial" phase conventions.[6] As human beings, there are ways to acquire an intuitive understanding of what a fictitious force is in the context of a non-inertial coordinate system. For instance, it is very difficult to keep one's balance when standing on a rotating platform because our senses are not accustomed to adjusting to the fact that the surroundings are not moving inertially. On a larger scale, we can observe phenomena that seem to violate Newton's laws of motion if we use the earth as an inertial reference frame. For instance, Buys Ballot's laws about wind directions cannot be understood without introducing the Coriolis force that is caused by the earth's rotation.

On the other hand, the notion of "phase" is one that is much further from everyday experience. Moreover, the first unified theory of electromagnetism, Maxwell theory, was already gauge invariant in the sense that the A_μ field (not expressed by Maxwell in this form) possessed "gauge-freedom". The notion of a non-inertial gauge convention was therefore not an issue. It was not before quantum mechanics, and the theory of quantum particles interacting with electromagnetic fields, that it was realized that the A_μ field has a role to play with respect to the phase. But even then, the impossibility of measuring the phase of a particle at a point (only phase differences can be measured in interference effects) makes the notion of a choice of phase convention seem artificial, or superfluous to the purposes of observation. This probably also accounts for why there is never any discussion of "fictitious phase effects". By this I mean phase-related behaviour of quantum particles that does not seem to follow the standard equations.

From these considerations I conclude that, abstracting from their role in the context of discovery and their different relations to our everyday experience, the equivalence principle and the gauge principle can be seen to play similar roles in the logical structure of both GR and gauge theory respectively. Both allow a certain mathematical structure, the Christoffel symbols $\Gamma^\lambda_{\mu\nu}$ and the (electromagnetic) potential A_μ, to be introduced into their respective theories, whose properties can then be generalized and extended to represent physical features of the world. In other words, both

[6] I put "non-inertial" in scare quotes because I have not yet properly developed an extension of the notion of inertia to phase. However, this will be done later.

principles make possible, in a similar way, the geometrization of gravitation and quantum-electromagnetism. I believe that this is the best answer that can be given to the mystery identified by Teller: "How can an apparently substantive conclusion follow from a fact about conventions?" [Teller, 2000, p. 469].

The final objection to the gauge principle is that it does not generate a field which satisfies the "action/reaction principle". In other words, Brown explains that the A_μ field would satisfy this principle when the "the matter field acts back on the connection [the A_μ field]" [Brown, 1999, p. 8]. He would like to see "the introduction of an analogue of Einstein's field equations, which would determine *inter alia* the effect of the Dirac particle on the gauge potential A_μ". In fact, in the case of a Dirac particle, the analogue of the Einstein field equations that Brown has in mind are the Maxwell equations for the electric and magnetic fields with the electric current density j^μ proportional to the current $\bar{\Psi}\gamma^\mu\Psi$. This form of the Maxwell equations can be obtained by varying, with respect to A_μ, the Dirac Lagrangian (4.15) with the Maxwell Lagrangian added

$$\mathcal{L}_{Dirac_{int}} + \mathcal{L}_{Maxwell} = \bar{\Psi}(i\gamma^\mu(\partial_\mu + iqA_\mu) - m)\Psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{4.21}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. In a mathematical form, Brown's objection, echoed by Martin, is that the gauge principle does not require the addition of $\mathcal{L}_{Maxwell} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ to the initial Dirac Lagrangian. This term, which "gives physical life to the field" is "ultimately put in by hand" [Martin, 2002, p. 229].

This objection raises interesting questions about the general form in which a theory should be expressed. General relativity was not born formulated in terms of a Lagrangian. Instead, one can understand the theory as a collection of equations and rules for using these equations to describe physical phenomena. In this way, we know that particles will move according to the geodesic equation, and that the terms which appear in this equation, such as the Christoffel symbols, are in turn given by other equations, such as (4.3), which in turn contain terms, such as the metric, which are determined by yet more equations, the Einstein field equations. Under such an interpretation, it would be unreasonable to expect the equivalence principle to somehow also generate the EFE's. Instead, the equivalence principle introduces terms into the theory whose roles can then be extended by the *deliberate addition* of new equations to the theory.

This reasoning can be applied *quid pro quo* to the gauge principle. Thus, we use the gauge principle to introduce A_μ into the equations of motion of our theory (such as the Schrödinger equation), and we then extend the role of this mathematical object by subjecting it to additional field equations, such as the Maxwell equations (for the electromagnetic potential). Given this notion of the form of a theory, the Lagrangian context seems like the wrong environment in which to judge the role of the principle.

As an aside, a closer look at the Lagrangian formulations of GR and quantum-electrodynamics (a fully quantized version of quantum electromagnetism) highlights important differences between the two theories. In a Lagrangian formulation of GR, the EFE's are obtained by varying the Einstein-Hilbert action [Misner et al., 1973, p. 486]

$$\mathcal{L}_{E-H} = \frac{-1}{2\kappa} \int R\sqrt{-g}\mathrm{d}^4 x \tag{4.22}$$

where $g = det(g_{\mu\nu})$ is the determinant of the metric tensor, R is the Ricci scalar and $\kappa = 8\pi Gc$, where G is Newton's gravitational constant and c is the speed of light. In quantum electrodynamics, one can start from the Maxwell Lagrangian and, given that $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ derive a propagator for the A_μ field. This is a crucial step towards giving a quantized theory of electromagnetism using the path integral approach. However, in GR the situation is more complicated. In general, the propagation of gravitational waves is studied in a linearized version of GR [Misner et al., 1973, p. 493]. It is well known that this approach does not succeed in allowing a quantization of the gravitational field. It has been suggested that difficulties in quantizing GR are related to its dependence on both a metric and a connection, whereas other gauge theories only rely on a connection. This has lead some to believe that attempts to quantize the gravitational field should focus on the Christoffel symbols, rather than on the metric [Anandan, 1993].

The final remark on the fourth objection is that it does not completely capture the role that an analogue to the Einstein field equations should play in Q-EM. As I will argue in the next section, there are additional field equations for A_μ that can be added to the theory and which are not derivable from the additional term $\mathcal{L}_{Maxwell}$. These additional equations, called the "holonomy equations", are motivated by the Aharonov-Bohm effect and are necessary to determine the values of A_μ that are relevant to its interaction with the phase of a quantum particle. Furthermore, I will argue that these holonomy equations are in a sense a better analogue to the Einstein field equations than Maxwell's equations. This furthers my conviction that the gauge principle should not be judged on its ability to precisely generate "complete" interaction Lagrangians.

4.3 Part II: Field equations and inertiality

One of the conclusions that can be drawn from the interpretation of the equivalence principle that I suggested in the previous section is that in general relativity, Einstein achieves a unification of the inertio-gravitational field. Mathematically, this means that both gravitational effects (linked to sources) and inertial effects (caused by the choice of a non-inertial coordinate system) are represented in the theory by the same mathematical object, the Christoffel symbols $\Gamma^{\lambda}_{\mu\nu}$. Since this object is not a tensor, its transformation properties [Weinberg, 1972, p. 100].

$$\Gamma'^{\lambda}_{\mu\nu} = \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial x^{\tau}}{\partial x'^{\mu}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \Gamma^{\rho}_{\tau\sigma} + \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial^2 x^{\rho}}{\partial x'^{\mu} \partial x'^{\nu}} \tag{4.23}$$

suggest that it is always possible to find a local coordinate system ξ^{α}_{X} (defined at a spacetime point X), such that the Christoffel symbols vanish, and the equations of motion for a test particle become those of a force free particle, as given by equation (4.1), repeated here

$$\frac{\mathrm{d}^2 \xi^{\alpha}_{X}}{\mathrm{d}\tau^2} = 0 \tag{4.24}$$

where τ is the proper time of the particle. Because the equations of motion take this form in it, the coordinate system ξ^{α}_{X} is often referred to as a local inertial coordinate system.[7] The Christoffel symbols are determined by a solution to the Einstein field equations (via the metric), whose "input" is the stress-energy tensor $T_{\mu\nu}$, which can be considered a contingent feature of the universe, analogous to a boundary condition. Since the Christoffel symbols determine the local inertial coordinate systems, Dieks argues that in GR, the property of being an inertial coordinate system is demoted from a *de jure* (law-like) feature to a *de facto* (contingent) one [Dieks, 2006, p. 15]. In other words, whether an arbitrary coordinate system at arbitrary spacetime point X, x^{μ}_{X}, is an inertial one cannot be determined *a priori*, in other words before having solved the EFE's.

The difference between the contingent and a priori inertiality of coordinate systems is best understood through a comparison with special relativity. Consider you are given a certain coordinate system x^{μ} on a Minkowski spacetime, and are asked to determine whether it is inertial or not. Using the

[7] Given the relationship between the metric and the Christoffel symbols, the local inertial coordinate system is one for which the first derivatives of the metric vanish. It is also interesting to note that this line of reasoning suggests how the "infinitesimal" principle of equivalence, never adhered to by Einstein, can nevertheless be seen as an outgrowth of his unification of the inertio-gravitational field.

Minkowsi metric, you could trace out an inertial worldline in the spacetime manifold and then look at how points along this worldline are coordinatized by x^μ. If the coordinates obey the inertial equation of motion (4.1), the x^μ is an inertial coordinate system. Now try repeating the same test in general relativity, before having solved the EFE's. One stumbles at the first hurdle: without a metric, it is impossible to trace out an inertial worldline in the spacetime manifold. Since in GR there are no "absolute objects" (in Anderson's sense) which can be appealed to before a solution of the EFE's is at hand, given an arbitrary curve in the manifold, there is *nothing* to distinguish a coordinatization of this curve by one or another coordinate system. One concludes that all coordinate systems must be equivalent.

From these considerations, I conclude that the Einstein field equations can be interpreted as defining a field (the metric), which breaks the symmetry of the infinitely many different possible coordinatizations of the spacetime manifold. Before solving the EFE's all the coordinate systems were equivalent, but afterwards, they can be distinguished by the values that they assign the Christoffel symbols. In the next section, I will propose a set of equations that play an exactly analogous role with respect to choices of gauge. In this way, we will be able to rigorously define the notion of an "inertial gauge".

4.3.1 Force-free motion in quantum-electromagnetism

In the previous section we argued that local inertial coordinate systems should be defined as those in which the laws of motion take a particularly simple form, namely when the Christoffel symbols vanish. Our analogy between the Christoffel symbols and the electromagnetic vector potential A_μ suggests that an inertial gauge should be defined by being the one in which the laws of motion take a particularly simple form, namely when the electromagnetic potential vanishes. The problem is finding the appropriate equation which must take this particularly simple form. In other words, what is the "something" whose laws of motion can be made particularly simple by choosing a particular choice of gauge?

The first obvious place to look is at the motion of massive *charged* particle, one that is affected by the presence of electromagnetic $\mathbf{E}$ and $\mathbf{B}$ fields. These fields contribute an additional force term, called the Lorentz force, to the equations of motions of a particle of charge q, moving with 3-velocity $\mathbf{v}$, which is given by

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{4.25}$$

In relativistic notation, this equation can be written in terms of the Minkowski force K^α [Griffiths, 1999, p. 540]

$$K^\alpha = \frac{\mathrm{d}p^\alpha}{\mathrm{d}\tau} = qU_\beta F^{\alpha\beta} \tag{4.26}$$

with p^α the 4-momentum of the particle, τ its proper time, U_β its 4-velocity and $F^{\alpha\beta}$ the electromagnetic tensor. Thus, in the absence of gravitation and in an inertial coordinate system, the equations of motion for this particle can be written as

$$m\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}\tau^2} = qU_\beta F^{\alpha\beta} \tag{4.27}$$

where m is the mass of the particle. Following the analogy with the equivalence principle, we might now try to define the "inertial gauge" as the one for which these equations return to their inertial form, in other words for which the right-hand side vanishes. Unfortunately, this is impossible, because $F^{\alpha\beta}$ is gauge-invariant, and therefore no gauge transformation can make the right-hand side vanish. I conclude that another kind of "motion", described by different equations, is needed to make possible the definition of an inertial gauge.

As we have already seen, another context in which the A_μ field plays a role is the Schrödinger equation. The gauge argument demonstrated that the introduction of a mathematical object, that we identified as a special zero-field version of the electromagnetic potential, was necessary to compensate for arbitrary phase conventions. Thus, it seems that in the absence of electromagnetic sources, we can define the inertial gauge as the one in which the Schrödinger equation takes its "standard" form (4.4) (repeated here)

$$\frac{-1}{2m}\nabla^2\psi(\mathbf{x},t) = i\frac{\partial\psi(\mathbf{x},t)}{\partial t} \tag{4.28}$$

In a non-inertial gauge, where the phase of the wavefunction differs per point according to $\psi' = e^{iq\chi(\mathbf{x},t)}\psi$, the Schrödinger equation takes the form (4.9) (repeated here)

$$\frac{-1}{2m}(\nabla - iq\mathbf{A})^2\psi'(\mathbf{x},t) + q\phi\psi'(\mathbf{x},t) = i\frac{\partial\psi'(\mathbf{x},t)}{\partial t} \tag{4.29}$$

where $A^\mu = ((-\frac{\partial}{\partial t}\chi(\mathbf{x},t), \nabla\chi(\mathbf{x},t))$ for some $\chi(\mathbf{x},t)$. This entails that it is possible to find a joint gauge transformation of the matter fields and electromagnetic potential which returns (4.29) to (4.28).

At this point, we are in a situation analogous to special relativity. We have defined the notion of an inertial gauge, but it seems to be determined by an absolute "phase-background". In other words, $\chi(\mathbf{x}, t) = constant$, plays the role of an absolute object in this theory. In any quantum mechanical situation we can imagine in which there are no electric and magnetic fields, the inertial gauges are ones for which the wavefunction can be written as $e^{iq\chi(\mathbf{x},t)}\psi(\mathbf{x}, t)$, with $\chi(\mathbf{x}, t)$ a constant. In order to remove this absolute object from the theory, it is necessary to make the gauge in which the equations take their simplest form the subject of a field equation, which must have only contingent features of the particular quantum mechanical situation under study as inputs. It was already suggested by Brown that the Maxwell equations could play this role. However, we have argued that the context in which the **E** and **B** fields become relevant is not conducive to defining the notion of inertiality for gauge. Instead, we need a quantum mechanical context in which the A_μ field is determined by *contingent* features of the situation.

The Aharonov-Bohm (AB) effect provides us with just such a context, in which the A_μ field plays an important role in determining the equations of motion for the phase of a quantum particle. The AB effect shows how the role of electromagnetic vector potential can be extended from correcting for non-standard phase conventions to encoding physical effects of electromagnetic sources on the phase of a quantum particle. This duality of the role of the vector potential in Q-EM (accounting for both non-inertial phase conventions and encoding the presence of physical electromagnetic influences) is exactly analogous to the dual role of the Christoffel symbols in GR.

4.3.2 The Aharonov-Bohm effect

In chapter 1 we cited a question, posed by Bohm in 1952, about the possibility of "a geometric or mechanical transformation" corresponding to a gauge transformation. His 1959 prediction, together with Aharonov, of what is now called the Aharonov-Bohm effect, can be seen as an answer to this question [Aharonov and Bohm, 1959]. The Aharonov-Bohm effect shows that electric charges and magnetic fluxes can be used to alter the phase of the wavefunction of a quantum particle, and thereby affect the interference pattern that such a particle can give rise to, for instance in a double slit setup. Most surprising of all, the "interaction" between the electromagnetic flux and the phase can take place even if the flux is shielded from the particle, in other words the wavefunction vanishes in the region where the flux is. This shows that knowledge of the electric and magnetic fields at each point where the wavefunction does not vanish is not sufficient to predict the state of a quantum mechanical system. In addition to the electric and magnetic fields, attention must be paid to global properties of the space in which the system evolves, encoded in the electromagnetic vector potential.[8] The integrability of the wavefunction, as determined by the vector potential, will have empirical consequences.

The Aharonov-Bohm effect comes in two variants: the electric effect and the magnetic effect [Peshkin and Tonomura, 1989]. Aharonov and Bohm's original paper first presents the electric version and then claims that "relativistic considerations" point towards a similar magnetic one [Aharonov and Bohm, 1959, p. 486]. However the magnetic effect can be more rigorously derived by appealing to the the condition that the wavefunction be singled valued. AB also present this derivation, and it is the one that is found most generally in the literature [Aharonov and Bohm, 1959, p. 486], [Peshkin and Tonomura, 1989]. I will start by the presenting the electric effect, before moving on to the magnetic one.

The electric AB effect

AB start by considering a charged quantum particle inside a Faraday cage (as discussed in chapter 1) described, in the absence of any charges on the cage, by the wavefunction $\Psi_0(\mathbf{x}, t)$ [Aharonov and Bohm, 1959, p. 485]. By adding charge to the surface of the cage, the potential inside the cage is altered uniformly without causing electric or magnetic fields to appear there. If the potential in the cage is varied over a time period $[t_0, t]$, then the phase of the quantum particle will be shifted by an amount (in natural units) e^{-iqS}, with $S = \int_{t_0}^{t} \phi(t')\mathrm{d}t'$. After performing this operation, the particle is desribed by the wavefunction $e^{-iqS}\Psi_0(\mathbf{x}, t)$. Note that this potential induced phase factor is achieved without resorting to any non-zero electric or magnetic fields in the region where the particle is located (the region in which the wavefunction of the particle doesn't vanish).

AB then consider taking a single coherent electron beam, splitting it into two beams, and passing each one through its own Faraday cage. This setup is illustrated in figure 4.2. The two beams will suffer a phase shift of e^{-iqS_1} and e^{-iqS_2} respectively, with $S_1 = \int_{t_0}^{t} \phi_1(t')\mathrm{d}t'$ and $S_2 = \int_{t_0}^{t} \phi_2(t')\mathrm{d}t'$. When the two beams are recombined and allowed to interfere, the interference pattern will depend on the difference between the phase shifts undergone by the two beams, $S_1 - S_2$. AB conclude that there is a "physical effect of the potentials" because the interference pattern has been changed while the wavefunction has only traversed regions in which the $\mathbf{E}$ and $\mathbf{B}$ fields vanished, but the

[8] One does not have to choose to encode the global properties in the vector potential. Any other method, for instance an appeal to the holonomies of the space, is also suitable. In what follows, I will not endorse any particular stance on the controversial issue of the "reality" of the vector potential. I will simply use it as a convenient tool for formulating certain predictions of Q-EM.

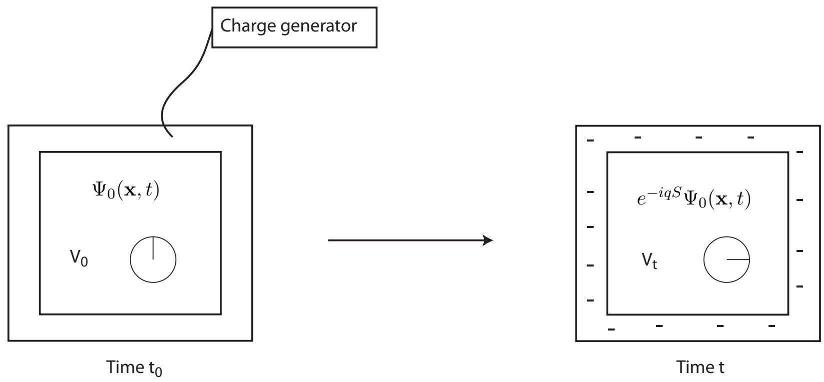

Figure 4.1: This figure shows the effect of charging a Faraday cage on the phase of a wavefunction contained inside the cage. The phase is represented by a clock-like figure. By charging the cage, and thus changing the potential inside the cage from V_0 to V_t, the phase of the wavefunction is rotated.

potentials did not.[9] They also note that this effect is "essentially quantum mechanical" [Aharonov and Bohm, 1959, p. 486]. This view is shared by Peshkin, who remarks that the "AB effect is deeply involved with the most primitive and general features of quantum theory" [Peshkin and Tonomura, 1989, p. 4].[10]

The magnetic AB effect

I will now derive the magnetic AB effect. As before, we take the zero-field solution to the Schrödinger equation to be $\Psi_0(\mathbf{x}, t)$. In the presence of a magnetic field, this solution becomes $e^{-qS}\Psi_0(\mathbf{x}, t)$ with $S = \int^x \mathbf{A}(x') \cdot \mathrm{d}\mathbf{x}'$, where x denotes the end point of integration [Peshkin and Tonomura, 1989,

[9] In chapter 5, we will argue that this is one way of implementing an active gauge transformation.

[10] I stress this point in order to put Wallace and Timpson's unorthodox interpretation of the AB effect into perspective. They claim that the AB effect is a "*classical* example of non-separability", where "classical" is to be understood as opposed to "quantum" [Wallace and Timpson, 2010, p. 714]. Their argument relies on the claim that such an effect is derivable in the context of the *classical* field theory of a complex field [Wallace, 2009]. Furthermore, they argue that such an effect would be testable experimentally if the classical limit of a complex scalar field were accessible to observation. However, I would object to where Wallace and Timpson draw the classical/quantum divide. One might want to argue that any considerations of a *complex* matter field already belong to the domain of quantum theory, even before this matter field is "quantized", in the sense that it is represented by an operator on Hilbert space.

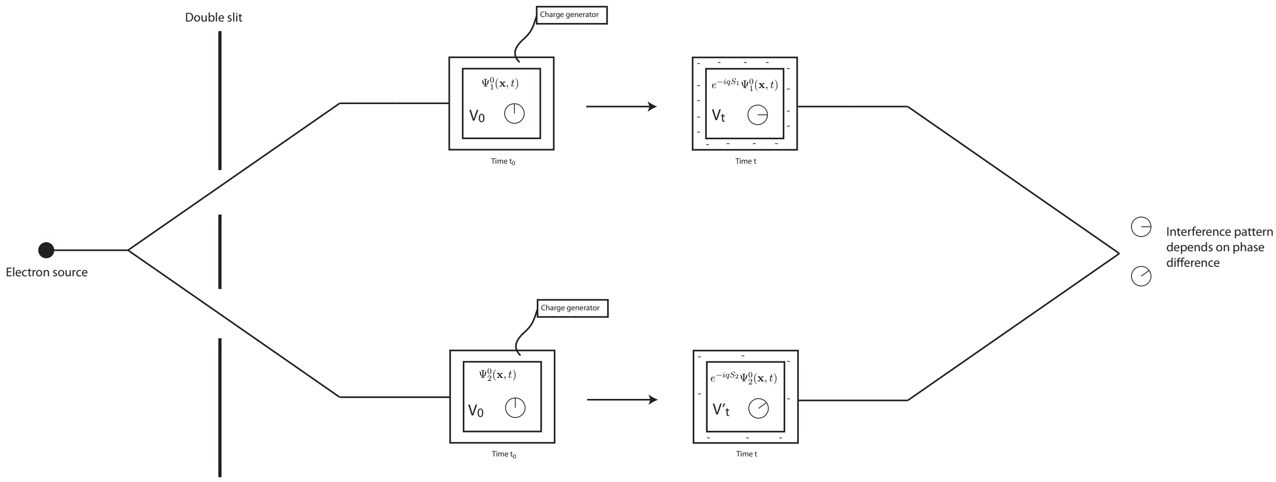

Figure 4.2: This figure depicts the electric AB effect. An electron beam passes through a double slit. Both the top beam and the bottom beam go through Faraday cage devices which are used to induce different phase-shifts in the wavefunctions of each beam. When the two beams recombine and interfere, the different phase-shifts will result in a shifted interference pattern when compared with that obtained in a normal double slit experiment.

p. 6].[11] Consider placing a thin solenoid between the slits through which a magnetic flux passes and calculating the value of the wavefunction at the point x on the screen where the interference pattern is measured. The setup is shown in figure 4.3.

If the integral is taken from the electron source to x along path 1, the solution for the wavefunction at x is $\Psi_1(\mathbf{x}, t) = e^{-qS_1}\Psi_0(\mathbf{x}, t)$ with $S_1 = \int_{\text{path1}} \mathbf{A}(x') \cdot \mathrm{d}\mathbf{x}'$. If the integral is taken along path 2, the solution for the wavefunction at x is $\Psi_2(\mathbf{x}, t) = e^{-qS_2}\Psi_0(\mathbf{x}, t)$ with $S_2 = \int_{\text{path2}} \mathbf{A}(x') \cdot \mathrm{d}\mathbf{x}'$. I will now show how the presence of the AB solenoid poses a problem for the requirement that the wavefunction be single-valued at x. Consider the loop integral around the coil, starting and ending at the source, traversing first path 1 and then path 2 in reverse.

[11]Technically, the start point of integration is $-\infty$ as in (4.20). However, since we will be considering loop integrals, the integration from $-\infty$ to the starting point can be discarded.

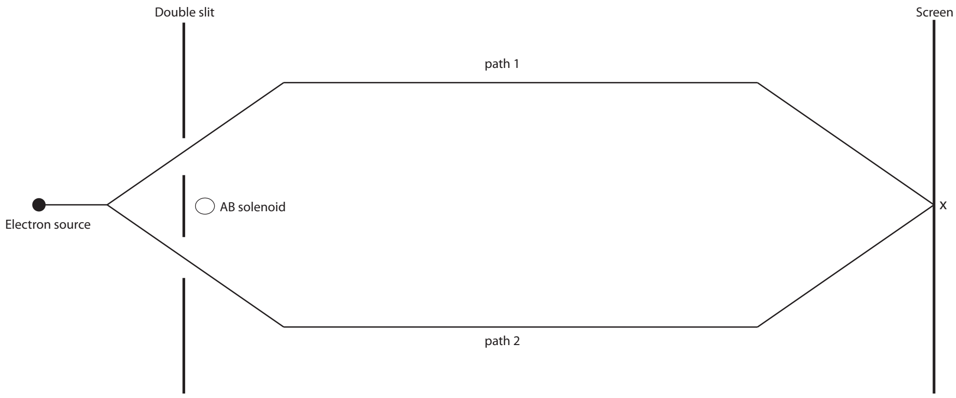

Figure 4.3: This figure shows the setup for the magnetic AB effect.

$$\begin{aligned}
&\int_{Source}^{x} \mathbf{A}(x')\cdot \mathrm{d}\mathbf{x}' - \int_{Source}^{x} \mathbf{A}(x')\cdot \mathrm{d}\mathbf{x}' \\
&= \int_{Source}^{x} \mathbf{A}(x')\cdot \mathrm{d}\mathbf{x}' + \int_{x}^{Source} \mathbf{A}(x')\cdot \mathrm{d}\mathbf{x}' \\
&= \oint \mathbf{A}(x')\cdot \mathrm{d}\mathbf{x}' \\
&= \int (\nabla \times \mathbf{A})\cdot \mathrm{d}\mathbf{S} = \int \mathbf{B}\cdot \mathrm{d}\mathbf{S} = \Phi
\end{aligned}$$

where the final line is obtained by applying Stokes' theorem, and Φ is the magnetic flux through the solenoid. If the magnetic flux does not vanish neither will the loop integral around the coil, and therefore we conclude that the value of the wavefunction at x appears to depend on the path taken between the source and the screen. The AB coil has therefore affected the integrability of the wavefunction. In order to maintain the single valuedness of the wavefunction, we conclude that the path dependent wavefunctions Ψ_1 and Ψ_2 will interfere with each other at point x, giving the final wavefunction $\Psi = \Psi_1 + \Psi_2$ [Aharonov and Bohm, 1959, p. 486]. Due to the phase shift between Ψ_1 and Ψ_2, the interference pattern observed at the screen will differ from that obtained in the normal double slit setup (with no AB solenoid). This completes the derivation of the magnetic AB effect. Experimental verification of the Aharonov-Bohm effect has been achieved, by, among others, Chambers (in 1960) and Tonomura et al (in 1986) (Chambers [1960], Tonomura et al. [1986]).

4.3.3 Field equations and the inertial gauge

I will now suggest "field-equations" for the electromagnetic potential which can be used to define the notion of an inertial gauge. The Aharonov-Bohm effect shows that a magnetic flux Φ can act as a source of the electromagnetic potential, in the context of the "phase-motion" of a quantum particle. The field equation that links source to potential is

$$\oint \mathbf{A} \cdot \mathrm{d}\mathbf{x} = \Phi \tag{4.30}$$

This implies that $\mathbf{A}$ can no longer be written in the "impotent" form $\mathbf{A} = \nabla\chi(\mathbf{x}, t)$. If it could, then the loop integral would have to vanish

$$\oint \mathbf{A} \cdot \mathrm{d}\mathbf{x} = \oint \nabla\chi(\mathbf{x}, t) \cdot \mathrm{d}\mathbf{x} = \int \nabla \times \nabla\chi(\mathbf{x}, t) \cdot \mathrm{d}\mathbf{S} = 0 \tag{4.31}$$

Since $\mathbf{A}$ cannot be written as the gradient of some smooth function of space and time, it is impossible to find a global gauge transformation (a gauge transformation defined on the whole of spacetime), such that the electromagnetic potential vanishes everywhere, and the Schrödinger equation takes its ordinary form (4.4). This non-integrability of the gauge field is an important and well-studied feature of gauge theories [Wu and Yang, 1975]. In such cases, the vector potential can only be transformed away *locally*, at a point x, by selecting a gauge in which, at x, the Schrödinger equation takes its ordinary form (4.4). If (4.30) does not vanish, then a choice of gauge that was inertial when it did vanish, may no longer be inertial afterwards. The notion of an "absolute" inertial gauge, or a background against which inertiality can be defined, has thus been eliminated from the theory. Inertiality is now a property of a gauge contingent on a solution to the holonomy equations.

In the fibre bundle formulation of quantum-electromagnetism, such integrals over closed loops are called "holonomies", and can be taken to represent global topological properties of the base space (see chapter 3). This has led to a topological interpretation of the A-B effect. The electromagnetic influence of the AB coil on the phase of a quantum particle is geometrized by encoding it in the topology of the base space. This is a different geometrization from the case of GR, where gravitational influences are encoded in the curvature of spacetime. Nounou argues that an appeal to topology can help explain how the AB solenoid can have a non-local effect on the quantum particle [Nounoun, 2003, p. 194]. While I do not doubt the mathematical efficacy of exploiting the topological properties of the base space to encode electromagnetic influences, I disagree with Nounou's claim that such a move has more explanatory strength than other approaches to the AB effect. While the non-local influence of the AB solenoid may be puzzling, an appeal to the topological properties of space is no less so. Nounou suggests that an intuitive grasp of the effect of the solenoid on the topology can be obtained by realising that, due to the solenoid "a very big chunk of space, 10 000 000 000 bigger than the electron itself cannot be accessed by it" [Nounoun, 2003, p. 191]. However, as Nounou herself notes, the fact that there is a region of space that is not accessible to the electron is not relevant to the occurrence of the AB effect. It is the fact that a magnetic flux is present in this region. The inaccessibility of a large part of space to the electron is therefore not relevant to developing an intuitive grasp of the effect of the solenoid on the topology.

Chapter 5

A classification of transformations

In this final chapter I present a classification of the transformations that can be performed in three different theoretical contexts: a) classical spacetime theories (Newtonian mechanics in Galilean spacetime, special relativity in Minkowski spacetime and general relativity), b) non-relativistic quantum mechanics and c) gauge theories (as applied to both classical and quantum systems). For each transformation, I will identify whether it is an active or a passive transformation. In the case of active symmetry transformations, I will discuss how the transformation in question passes the "transformation condition" (introduced in chapter 1) [Brading and Brown, 2004, p. 646]. For the passive transformations I will introduce a distinction between weak and strong symmetries. This distinction will be relevant to understanding when passive symmetry transformations imply corresponding active symmetries and when this implication fails. In the case of classical spacetime theories and quantum mechanics, the classification that I propose follows orthodox views. However, in the case of the gauge theories, some ideas will go against the currently accepted interpretation, and I will address existing objections.[1] Finally I will reflect on the empirical significance of active transformations. I will conclude that the slightly different analysis that is necessary to make sense of active gauge transformations on quantum systems is due to fundamental differences between such systems and classical systems.

[1]An as yet unpublished pre-print by Wallace and Greaves presents ideas that are closer to mine than can be found elsewhere [Wallace and Greaves, 2011]. However, I disagree with some of their conclusions, as will be explained.

5.1 Tools of classification

Before beginning the classification of transformations proper, I introduce the tools that I will use for this classification.

5.1.1 Active and passive transformations

The definitions of active and passive transformations were given at the end of chapter 1. The crux of the distinction is that a passive transformation relates two different representations of the *same* physical situation whereas an active transformation relates two representations of *different* physical situations. In order to assess whether a given transformation is active or passive, it is therefore necessary to relate the representations connected by the transformations to actual physical situations. Before discussing in more depth how my use of the terms "active" and "passive" relates to their standard use in the literature, I clarify the notions of a "physical situation" and a "physical difference".

Physical situations and physical differences

By "physical situation" I mean a situation in the world. In other words, I mean the events that happen "out-there". By mathematical situation, I mean a representation of a physical situation. A physical situation is therefore a piece of "reality", whereas a mathematical situation is a piece of mathematics. The distinction between empirical and theoretical symmetries introduced in chapter 1 was intended to highlight the distance that exists between transformations that act on physical situations, and transformations that act on mathematical situations. Theories in physics that adequately represent the physical situations for which they are intended should be finely tuned enough that any change in the physical situation is captured by a change in the mathematical situation given by the theory. However, the converse is not necessarily true. It may be that a change in the mathematical situation does not correspondence to any change in the physical situation. We called such changes passive transformations. In Redhead's terms, a theory that gives rise to mathematical situations that are related by passive transformations has "surplus structure". However, a change in the mathematical situation can also correspond to a change in the physical situation it represents. Such a change is called an active transformation.

In distinguishing between active and passive transformations, clear criteria must be given for when the differences between mathematical situations correspond to differences between the physical situations they represent. For

Brown and Brading, the relevant criteria is "empirical significance" [Brading and Brown, 2004, p. 646]. A transformation between mathematical situations has empirical significance if there is an empirical difference between the physical situations that they represent. By "empirical difference", I believe they intend a difference in the outcome of some imaginable experiment.

Grounding the notion of physical difference in imaginable measurement outcomes is perhaps not entirely satisfactory. Consider the case of the velocity of the universe in Newtonian mechanics. The value of this velocity is one that distinguishes mathematical situations allowed by the theory. It is generally accepted however that it does *not* distinguish physical situations, principally because the value of this velocity would not affect the outcomes of any imaginable experiments. In order to make this line of reasoning complete, a solid argument is necessary for why the velocity of the universe should not affect any measurement outcomes. I am unaware of any such definitive argument.

Another situation in which such a situation arises is the hole argument in general relativity. However, in this case Einstein's point coincidence argument provides a premise from which it follows that no imaginable experiment will be able to distinguish between two diffeomorphically related models $\langle \mathcal{M}, g_{ab}, T_{ab} \rangle$ and $\langle \mathcal{M}, d^* g_{ab}, d^* T_{ab} \rangle$. The point coincidence argument states [Einstein, 1997, p. 153]

> All our spacetime verifications invariably amount to a determination of spacetime coincidences. If, for example, events consisted merely in the motion of material points, then ultimately nothing would be observable but the meetings of two or more of these points. Moreover, the results of our measurings are nothing but verifications of such meetings of the material points of our measuring instruments with other material points, coincidences between the hands of a clock and points on the clock dial, and observed point-events happening at the same place at the same time.

Since the coincidences of events are the same in both models $\langle \mathcal{M}, g_{ab}, T_{ab} \rangle$ and $\langle \mathcal{M}, d^* g_{ab}, d^* T_{ab} \rangle$, it follows that no measurement will be able to distinguish them. Such a conclusion obviously stands and falls with the point coincidence argument, but at least there is a direct line of reasoning from possible measurement outcomes to the empirical significance of transformations between mathematical situations.

To summarize: a physical difference between mathematical situations is a difference in the outcomes of measurements that could be performed on the physical situations represented by the mathematical situations.

Active and passive transformations in the literature

In the literature, the terms "active transformation" and "passive transformation" are often used in the way I suggest [Brading and Brown, 2004], [Brading and Castellani, 2003, p. 1343]. However, the terms are also often used, even by the same authors, in a slightly different way. Although this different usage overlaps considerably with mine, some fundamental differences may cause much confusion if not isolated and clarified at this point. In GR, coordinate transformations are often referred to as passive, whereas differentiable maps of the manifold onto itself (point transformations or diffeomorphisms) are often referred to as "active". However, this terminology does not imply that point transformations always relate different physical situations. As Stachel notes [Stachel, 1993, p. 133]

> A point transformation, as an active transformation, is of potential physical significance since, as we will see, under certain circumstances it can be used to turn one physical model in another.

When used in this way, "active transformations" only *potentially* relate different physical situations. As we just mentioned, the physical indistinguishability of $\langle \mathcal{M}, g_{ab}, T_{ab} \rangle$ and $\langle \mathcal{M}, d^* g_{ab}, d^* T_{ab} \rangle$ leads to the pitfalls of the hole argument if it is maintained that these two mathematical situations are different in some physically significant sense. These pitfalls are avoided by accepting that such diffeomorphically related models represent the same physical situation. Brading and Brown quote Anandan's claim that this "solution to the hole argument abolishes the distinction between active and passive transformations" [Brown and Brading, 2002, p. 28]. Of course, what Anandan means here is that the difference between *mathematically* active and passive transformations is abolished, not the difference between active and passive transformations as I have defined them. In fact, I will argue later that, even accepting this solution to the hole argument, "mathematically active transformations", in other words point transformations, can still relate physically different situations under certain conditions.

In summary, in the literature, the concepts "active transformation" and "point transformation" are often used synonymously. I believe that these two terms should be kept apart, and that the term 'active transformation' should be reserved for describing a transformation that makes a physical difference, or, in Brown and Brading's terminology, has empirical significance. In this way, whether a point transformation is an active transformation is not something that follows analytically from their definitions, it is something which must be shown.

5.1.2 Global and local transformations

When discussing transformations and symmetries, another common but confusing distinction is that between *local* and *global* transformations. In general, transformations are defined in terms of a function on (a region of) spacetime. If this function is a constant, so that it can be treated as a parameter, then the transformation is usually called global. On the other hand, if the function is not constant, then the transformation is usually called local. These definitions imply that global transformations are a special case of the local ones. However, this is a very confusing situation because it is often claimed that global transformations are symmetries of a theory but that local transformations are not. If global transformations are a subclass of the local ones, then the previous statement is a contradiction. For the purposes of my classification, I adopt a more intuitive notion of the distinction between global and local. I will not take the function in terms of which the transformation is defined as determining whether the transformation is local or global. Instead, I will look at the effect of the transformation on the system in question. If the transformation affects all parts of the system uniformly, I will call the transformation global. If it affects some parts of the system differently from others, I will call it local. The significance of this different approach to the local/global distinction will be especially important in the case of gauge transformations.

A second difficulty relating to the global/local distinction will arise in our discussion of active transformations. In what follows I will want to consider both global and local active transformations. However, it might initially seem like active global symmetry transformations are impossible because they cannot relate physically different situations.[2] A global active transformation will affect everything in the universe in the same way. Furthermore, since it is a symmetry, the transformed and untransformed situations will be indistinguishable. Giving the universe a uniform velocity in Newtonian spacetime was an example of such a transformation. We concluded that such a transformation did not lead to a physically different situation and therefore that it was not an active transformation. It seems that such a line of reasoning should apply to any global active symmetry transformation.

The only way to salvage the notion of an active global symmetry transformation is to define it in terms of "isolated subsystems" of the universe [Brading and Brown, 2004, p. 646]. A global transformation means that the same transformation is performed across the whole subsystem, while the rest of the universe is left unchanged. Strictly speaking, such a transformation is

[2]There is no problem for global transformations that are not symmetries because the very fact that they are not symmetries can be used to differentiate the transformed and untransformed situations.

local (it does not affect all parts of the universe in the same way). However, it is important to accept that if one is going to make any sense of active global symmetries, one must accept to treat isolated subsytems essentially as their own self-contained universes. Transformations acting on the subsystem are then classified as global or local with respect to that subsystem and not the universe as a whole.

5.1.3 Weak and strong symmetries

In chapter 1, I introduced some of the difficulties involved in identifying when a particular transformation qualified as a symmetry. I showed that different plausible definitions of the concept of symmetry did not result in the same end classifications. In order to capture these ambiguities, I will define two different ways in which a passive transformation can be a symmetry, which I term *weak symmetry* and *strong symmetry*. While it will be difficult to give sharp criteria to distinguish these two notions, I believe that something like this distinction plays a crucial role in determining when passive transformations imply active equivalents and when they do not. I will now try to motivate the distinction, and I hope that its usefulness will become clearer when I apply it in context.

In this chapter, we are concerned with transformations that act on models of a theory. As defined in chapter 1, a model of a theory is a mathematical situation, consisting of variables and functions of these variables, which obeys the laws (equations) of the theory. Broadly speaking, a transformation acting on this model is a symmetry if the mathematical situation that it generates also obeys the equations of the theory, and therefore is also a model of the theory. This agrees with the definition of symmetry endorsed by Brown and Brading [Brading and Brown, 2004, p. 645]. However, we also saw in chapter 1 that this definition still allows room for maneuver because it is reliant on exactly how the equations of the theory in question are formulated. In other words, the syntax used to write the equations of a theory can play a role in determining which transformations are symmetries and which are not [Norton, 2003]. As is well known, this issue is at the heart of debates about the significance of general covariance in general relativity. A clever formulation of the laws will make them flexible enough to allow a larger number of transformations to count as symmetries than a less clever formulation.

The notion of a 'clever formulation of the laws' is best substantiated by comparing neo-Newtonian mechanics (Newtonian mechanics without a notion of absolute rest) to general relativity.[3] In force free neo-Newtonian mechanics,

[3]The comparison could just as well have been between special relativity and general relativity. Thus everything I say about neo-Newtonian mechanics in the rest of this section

the equation that determines whether a mathematical situation is a model of the theory is

$$\frac{\mathrm{d}^2\mathbf{x}(t)}{\mathrm{d}t^2} = 0 \tag{5.1}$$

with the vector $\mathbf{x}(t)$ denoting the spatial coordinates of a particle and t absolute Newtonian time.[4] Its equivalent in general relativity is the geodesic equation

$$\frac{\mathrm{d}^2x^\lambda}{\mathrm{d}s^2} + \Gamma^\lambda_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\frac{\mathrm{d}x^\nu}{\mathrm{d}s} = 0 \tag{5.2}$$

where the four-vector x^λ denotes the spacetime coordinates of the particle and s is a parameter along the worldline of the particle. In neo-Newtonian mechanics, a coordinate transformation results in a transformation of the spatial coordinates of the particle from $\mathbf{x}(t)$ to some new spatial coordinates $\mathbf{x}'(t)$. This transformation is a symmetry if $\mathbf{x}'(t)$ obeys (5.1). In the case of GR, a coordinate transformation results in a transformation from the coordinates x^λ to new coordinates x'^λ. However, in order to judge whether this transformation is a symmetry, one does not test x'^λ in (5.2). Strictly speaking, one first transforms (5.2) by adjusting the Christoffel symbols $\Gamma^\lambda_{\mu\nu}$ to the new coordinate system and then one plugs in the new coordinate values. (5.1) is invariant under coordinate transformations because the exact same equation is used to judge whether any given situation is a model of neo-Newtonian mechanics. On the other hand, it is an illusion in a sense that (5.2) is invariant under coordinate transformations because the components of the Christoffel symbols have to be adapted to the particular coordinate system that is being used before the equation can be applied.

This difference between the invariance of the equations of neo-Newtonian mechanics and the invariance of the laws of GR has been pointed to numerous times in the literature [Norton, 1993, p. 833]. However, it is not straightforward to formulate the difference in a totally unambiguous manner. One possibility is to consider the sense in which different coordinate systems can be considered equivalent in a particular theory. In the case of neo-Newtonian mechanics, consider the trajectory of a particle described by the coordinates $\mathbf{x}(t)$. In a different coordinate system, the same trajectory

could be applied to special relativity, adjusting the equations to account for the fact that coordinates transform according to Lorentz transformations, rather than Galilean transformations.

[4]When we say a mathematical situation is a model of a theory, the mathematical situation represents the history of a system, not the instantaneous state of the system. Equation (5.1) is then taken to hold throughout the history of the system [Brading and Castellani, 2003, p. 1344].

will be described by the coordinates $\mathbf{x}'(t)$. If both $\mathbf{x}(t)$ and $\mathbf{x}'(t)$ satisfy (5.1), then they are both models of neo-Newtonian mechanics. Furthermore, one could call $\mathbf{x}(t)$ and $\mathbf{x}'(t)$ *equivalent* descriptions because there is nothing in the equation (5.1) to distinguish them. On the other hand, if the motion of a particle in GR is described by two coordinates systems x^λ and x'^λ, then there is a possibility that they can be distinguished, even if they both obey the generally covariant equation (5.2). If a transformation of the Christoffel symbols is needed when changing between these coordinate systems, then the values of the components of the Christoffel symbols can be used to differentiate the two coordinate systems and they cannot be called equivalent. As Dieks puts it, "the numerical values of the coefficients $\Gamma^\lambda_{\mu\nu}$ that occur in the 'generally valid form' of the equation of motion [...] are quantities that encode the acceleration of the frame [coordinate system] that is used" [Dieks, 2006, p. 12]. In a non-generally covariant formulation of neo-Newtonian mechanics, as is used here, symmetry transformations always relate equivalent coordinate systems. On the other hand, in GR, symmetry transformations sometimes relate inequivalent coordinate systems. I call symmetry transformations that relate equivalent coordinate systems *strong symmetries*, and those that relate inequivalent coordinate systems *weak symmetries*.

5.2 The classification

I now apply the tools developed above to classify the possible transformations of classical spacetime theories, quantum theory and gauge theory.

5.2.1 Classical spacetime theories

In chapter 1, we saw that the transformation

$$\mathbf{x}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} \rightarrow \mathbf{x}'(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix} = \begin{pmatrix} x(t) + v_x t \\ y(t) \\ z(t) \end{pmatrix} \tag{5.3}$$

had a dual interpretation in terms of passive and active versions. This duality can be captured geometrically, when the theory is formulated on a 4-dimensional manifold, in terms of the distinction between coordinate transformations and point transformations. Coordinate transformations were defined in section 3.1.2. For two coordinate systems (U_1, ϕ_1) and (U_2, ϕ_2), a coordinate transformation is the mapping of $\mathbb{R}^4$ onto itself given by $\phi_2 \circ \phi_1^{-1}$. There are many different kinds of coordinate transformation. For instance, it is possible to change from Cartesian coordinates to polar coordinates.

While such a transformation may greatly simplify calculations, it has no particular consequences and will not interest us further. It is also possible to change the properties of a coordinate system, for instance its state of motion, or the position of its origin. These kinds of transformation can have important consequences. A point transformation is a mapping of the manifold $\mathcal{M}$ onto itself. I will only be interested in differentiable point transformations, therefore in what follows I will assume the expressions "point transformation" and "diffeomorphism" to be synonymous. Torretti notes that for each coordinate transformation $\phi_2 \circ \phi_1^{-1}$, there is a corresponding point transformation given by $\phi_2^{-1} \circ \phi_1$ [Torretti, 1983, p. 25 and p. 31].

In the context of a spacetime theory, we assume that the state of a (classical) system is given by a configuration of non-intersecting time-like curves in the manifold, where each curve is the worldline of a part of the system. All coordinate transformations are clearly passive transformations, because they only affect how the worldlines of a system are described in terms of 4-tuples of real numbers and involve no physical interaction with the system. On the other hand, a point transformation will change the worldlines of the system in the manifold, and is therefore a candidate for an active transformation. Whether a point transformation is an active transformation will depend on whether it can be shown to map one physical situation onto a different physical situation.

Neo-Newtonian mechanics

I will now classify the coordinate transformations and point transformations of neo-Newtonian mechanics. It is helpful to distinguish between the coordinate transformations corresponding to elements of the proper orthochronous Galilei group and the rest.[5] As we know from Wigner's philosophy of symmetry, the transformations corresponding to the elements of the Galilei group are symmetries of neo-Newtonian mechanics. These are time translations, spatial translations, rotations and boosts. By boost we mean that we transform from one coordinate system to another in a state of uniform motion with respect to the first (the passive version of (5.3)). Coordinate transformations which do not correspond to elements of the Galilei group are not symmetries. For instance, (5.1) is not valid under a transformation to a rotating coordinate system. In that case, additional terms accounting for the fictitious Coriolis and centripetal forces are needed to describe the motion of a system. The transformations corresponding to the elements of the

[5] I will often refer to the proper orthochronous Galilei group simply as the Galilei group. The time reversal and parity transformations will have no part to play in what follows. Although to some extent the Galilei group and the Poincaré group contain the same transformations, namely spatial translations, time translations, rotations and boosts, differences in the way these are defined mean that both groups have different invariants.

Galilei group are global because each point of the coordinate system suffers the same transformation (all the points are shifted by the same amount in the case of a spatial translation for example).

Consider now the point transformation $\phi_2^{-1} \circ \phi_1$ corresponding to the coordinate transformation $\phi_2 \circ \phi_1^{-1}$. I will show that this represents a physically different situation by applying the point transformation to an isolated subsystem. Let a system S be represented by a configuration of time-like worldlines in a region U of the manifold. In another region U' of the manifold, we define an environment system E, consisting of one time-like worldline. After applying the point transformation in the region U (and the identity transformation on the region U'), we can study the new relations between the environment and the transformed system S'. For instance, we can find the distance from E to a point of S' at a certain time t. After the point transformation, this distance will be different than it was before. This corresponds to a physical consequence of the point transformation, and thus we can conclude that the point transformation is an active transformation. Any point transformation that differs from the identity in the region U will result in a change in the relations between S and E. As we saw at the beginning of this chapter, the coordinate transformations corresponding to the elements of the Galilei group are strong symmetries, because they relate equivalent coordinate systems. This can be seen from the fact that it is not possible to use (5.1) to distinguish between two coordinate systems related by such a transformation. Let ϕ_1 and ϕ_2 be two such coordinate systems. It can be shown that if a worldline W is a model of neo-Newtonian mechanics, then $\phi_2^{-1} \circ \phi_1(W)$ is also a model [Torretti, 1983, p. 65]. In other words, the passive coordinate symmetry transformation implies the active point symmetry transformation.

Proof (passive symmetry implies active symmetry): Take a physical system S, represented by a configuration of timelike worldlines, determined by laws L and initial conditions $C(S)$. A point transformation $\tau = \phi_2^{-1} \circ \phi_1$ takes S into τS, and $C(S)$ into $C(\tau S)$. If ϕ_1 and ϕ_2 are related by a transformation corresponding to an element of the Galilei group, then the laws L will be the same in both coordinate systems. Note that $\phi_2(\tau S) = \phi_2(\phi_2^{-1} \circ \phi_1 S) = \phi_1(S)$. This means that the coordinates of the system S in ϕ_1 are exactly the same 4-tuples of real numbers as the coordinates of the system τS in ϕ_2. But since the same laws L apply in ϕ_1 and ϕ_2 and S obeys the laws, then τS will also obey the laws.

Note that the implication can also be proven in the other direction, making the passive and active versions of the symmetry logically equivalent. Torretti notes that Einstein also appreciated this logical equivalence of active and passive symmetries [Torretti, 1983, p. 65].

Proof (active symmetry implies passive symmetry): Take two physical systems S and S', represented by a configuration of timelike worldlines, such that $S' = \tau S$, and τ is a symmetry of neo-Newtonian mechanics. This means that both S and S' are models of neo-Newtonian mechanics. Choose a coordinate system ϕ_1 such that the coordinates of $\phi_1(S)$ obey the laws L of neo-Newtonian mechanics. Now choose a second coordinate system ϕ_2 such that $\phi_2(S') = \phi_1(S)$. Since S and S' are both models of neo-Newtonian mechanics, and $\phi_2(S') = \phi_1(S)$, the laws L must be the same in both coordinate systems. ϕ_1 and ϕ_2 are thus equivalent coordinate systems.

The point transformations corresponding to the elements of the Galilei group can be made empirically significant by applying them to an isolated subsystem of the universe. In this way, the above arguments show that all such point transformations are active symmetry transformations. Consider the point transformation that takes the worldlines of a system S into the worldlines of a system S' such that, in the inertial coordinate system ϕ_1, the system S' is in a state of uniform motion with respect to the untransformed system S. This is an active symmetry transformation that corresponds to the empirical symmetry of Galileo's ship, as described in chapter 1.

We mentioned earlier that the passive transformation from an inertial coordinate system to a rotating coordinate system was not a symmetry of neo-Newtonian mechanics as it is formulated here. The active point transformation corresponding to this transformation is also not a symmetry of force-free neo-Newtonian mechanics because the transformed situation does not obey the laws of the theory. However, it is possible to give a generally covariant formulation of neo-Newtonian mechanics in which the passive transformation to a rotating coordinate system becomes a symmetry. However, such a formulation would have to introduce elements into the equations which could be used to distinguish between inertial coordinate systems and rotating ones. In this case, the passive coordinate transformation from an inertial to a rotating coordinate system would be a weak symmetry and would not imply that the corresponding active point transformation became a symmetry. This analysis reveals why one should be cautious when claiming that generally covariant techniques extend the symmetry groups of a theory. In fact, a general covariant formulation can be said to extend the passive symmetries of a theory (in a sense), but it does not extend the active ones.

The following table presents the results of the classification for neo-Newtonian mechanics, as formulated in a non-generally covariant fashion.

	Active/Passive	Global/Local	Symmetry
Coordinate transformations (Galilei group)	Passive	Global	Yes (strong)
Coordinate transformations (other)	Passive	Both	No
Point transformations (Galilei group)	Active	Global	Yes
Point transformations (other)	Active	Both	No

Special Relativity

The classification of transformations in special relativity is identical to that of neo-Newtonian mechanics, only the details of the mathematical transformations change to take account of the structure of Minkowski spacetime, which differs from that of neo-Newtonian spacetime. The following Lorentz transformation (in natural units) is the special relativistic analogue of (5.3), and it also has a dual interpretation in terms of active and passive versions.

$$x^\lambda = \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} \to x'^\lambda = \begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} \tag{5.4}$$

with $\gamma = \frac{1}{\sqrt{1-v^2}}$ the usual Lorentz factor. The worldline W of a massive particle is a model of the theory if it is timelike and its length L given by

$$L = \int_W \mathrm{d}s \tag{5.5}$$

with the interval $ds^2 = \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$, is extremal. This is the case if

$$\delta L = \delta \int_W ds = 0 \tag{5.6}$$

As in the case of neo-Newtonian spacetime, it is not possible to use these equations to differentiate between two inertial coordinate systems. Let W be the worldline of a massive particle, and ϕ_1 and ϕ_2 be two inertial coordinate systems. Exactly the same equations will be used in both coordinate systems to judge whether W is a model of SR. More specifically, both coordinate systems will make use of the Minkowski metric $\eta_{\mu\nu}$ to calculate distances. This means that in SR, all inertial coordinate systems are equivalent, and therefore that transformations between them are strong symmetries. All coordinate transformations corresponding to elements of the Poincaré group are strong symmetries of SR, and they imply equivalent active symmetry transformations (which take the form of point transformations) [Torretti, 1983, p. 65]. The classification table for special relativity looks exactly the same as that for neo-Newtonian mechanics, with the Galilei group now replaced by the Poincaré group.[6]

Before moving on to general relativity, I will make a small remark about the equivalence of active and passive symmetry transformations that holds in neo-Newtonian mechanics and special relativity. The logical equivalence of active and passive transformations suggests that one cannot consider the one or the other as being more or less revelatory about the structure of our world. This seems to be in tension with Wigner's philosophy of symmetry as presented in chapter 2. We saw that Wigner believed that invariance of the laws of nature under a change of description (a passive symmetry) was a trivial matter, and therefore that only invariance after *physical* action on a system (an active transformation) could be of interest. In other writings, Wigner has recognized the significance of Einstein's use of symmetry principles in physics [Wigner, 1991]. However, a close look at the Relativity Principle of Einstein's 1905 paper shows that it is formulated in terms of passive transformations, precisely the kind which Wigner dismisses as trivial. The Relativity principle of 1905 states [Einstein, 1923, p. 4]

> The laws by which states of physical systems undergo change are not affected, whether these changes of state be referred to one or the other of two systems of coordinates in uniform translatory motion.

The apparent incoherence of Wigner's views can be resolved by pointing out an equivocation over what is meant by "laws of nature". Torretti explains [Torretti, 1983, p. 54]

> The "laws" mentioned by the RP [Relativity Principle] are not the real relations between classes of events that a philosopher would call by that name, but rather the relations between coordinate functions, and functions of such functions, by which the physicist seeks to express the former. If the RP spoke about real relations of events it would be trivial: obviously, such relations cannot be affected by the choice men make of a coordinate

[6]Given that the transition from neo-Newtonian mechanics to special relativity was supposedly revolutionary, it may initially seem odd that my classification of symmetries does not bring out any significant difference between them. However, as far as symmetries are concerned, special relativity does not add anything to neo-Newtonian mechanics, as the famous Relativity Principle was in fact already well known since Galileo. The revolutionary new concepts of space and time that emerge from SR come from combining the Relativity Principle with the Light Principle. This has very significant consequences for how one transforms from one coordinate system to another, which is reflected in the shift from the Galilei group to the Poincaré group. However, it does not affect which of these transformations are symmetries, and it does not affect which active transformations are symmetries.

system for describing them.

One must sharply distinguish between "laws of nature" understood as abstract, observer independent facts about how systems evolve, and "laws of physics" understood as equations written down on paper by physicists. The latter are not observer independent facts, because they must relate measurable features of systems, and measurements require observers and reference frames. In dismissing the significance of passive transformations, it seems that Wigner had the laws of nature in mind, rather than the laws of physics. In section 2.1.3, we described this position of Wigner's by saying that he often thought of the evolution of a system as an "evolution in-itself", and not as one written out in terms of some (mathematical) representation by an observer.

General Relativity

The dynamical nature of spacetime in GR makes it impossible to give a generally valid classification of its transformations. As in the case of SR, in force-free GR allowed worldlines of a massive particle are time-like and extremal.[7] They must therefore also obey (5.6). However the interval is now given by $\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$ where the metric $g_{\mu\nu}$ is a solution to the Einstein field equations (EFE's).

The general covariant formalism that is used implies that all coordinate transformations are at least weak symmetries. However, the coordinate transformations that are also strong symmetries will depend on the particular solution of the Einstein field equations at hand. For example, the Minkowski metric is a vacuum solution to the EFE's, and, as we have seen, admits the transformations of the Poincaré group as strong symmetries. However, it is also possible to imagine very uneven metrics which have no strong symmetries. In this case, every coordinate system can be identified by the values of the Christoffel symbols that it takes.

As already mentioned at the beginning of this chapter, learning the lesson taught by the hole argument implies that not all point transformations in GR can be considered as active transformations. However, the transformations considered in the hole argument are more than just point transformations because they also involve "dragging-along" the tensors defined on the manifold. Consider the structure $\langle \mathcal{M}, g_{ab}, T_{ab} \rangle$, with $\mathcal{M}$ a four-dimensional differentiable manifold, and g_{ab} and T_{ab} geometric objects representing the metric and the stress-energy tensor respectively. This structure is a model of GR in the sense that g_{ab} and T_{ab} satisfy the EFE's. Consider also the point

[7]The term 'force-free' is not meant to exclude curved-spacetimes, but only those additional forces that might come from electromagnetic or other (non-gravitational) effects.

transformation (diffeomorphism) $d : \mathcal{M} \to \mathcal{M}$. This point transformation can be used to define the dragged-along metric and stress-energy tensors d^*g_{ab} and d^*T_{ab}. It can be shown that the structure $\langle \mathcal{M}, d^*g_{ab}, d^*T_{ab} \rangle$ is also a model of GR [Norton, 1993, p. 824]. Furthermore, no experiments can be performed to distinguish $\langle \mathcal{M}, g_{ab}, T_{ab} \rangle$ from $\langle \mathcal{M}, d^*g_{ab}, d^*T_{ab} \rangle$. However, as Norton stresses, these two structures are "mathematically independent", in the sense that an event that happens at a certain point of the manifold in the first model is mapped by d onto a different point of the manifold in the second model. The hole argument can be used to show that if one interprets the initial structure and its transform as representing physically different situations, then general relativity becomes an indeterministic theory. Restoring determinism is achieved at the expense of interpreting $\langle \mathcal{M}, g_{ab}, T_{ab} \rangle$ and $\langle \mathcal{M}, d^*g_{ab}, d^*T_{ab} \rangle$ as representing the same physical situation. Accepting this solution to hole argument, I classify point transformations accompanied by drag-alongs of the tensor fields as passive symmetries of the Einstein field equations. I suspend judgment about whether they should be considered strong or weak symmetries. In fact, Earman and Norton have suggested that point transformations that drag along the tensor fields on the manifold can be defined in neo-Newtonian mechanics and special relativity as well, provided that these theories are reformulated in a certain way [Earman and Norton, 1987]. In these cases, an analogue to the hole argument is also possible. My analysis suggests that in all these cases, the point transformations that also drag-along the tensor fields should be considered as passive transformations of the relevant field equations.[8]

There have been attempts to rescue the deterministic character of GR while maintaining that $\langle \mathcal{M}, g_{ab}, T_{ab} \rangle$ and $\langle \mathcal{M}, d^*g_{ab}, d^*T_{ab} \rangle$ represent different physical situations. However, such a move is not necessary to preserve the claim that point transformations can, under certain conditions, be interpreted as active transformations in GR. A crucial feature of the hole argument is that the point transformation is also used to transform the metric and the stress-energy tensors. If one considers point transformations that act within a solution to the EFE's (without dragging-along the metric and the stress-energy tensor), then they can be used to generate physically different situations just as they can in Galilean and Minkowski spacetimes. Within a solution of the EFE's a local point transformation or a global point transformation that acts on an isolated subsystem will change the relations that hold between the particles evolving in the spacetime. These changes will be empirically detectable. From these considerations, I conclude that point transformations associated with dragged-along tensor fields are pas-

[8] In such reformulated versions of neo-Newtonian mechanics and special relativity, not all point transformations are therefore active transformations. The extent to which it is reasonable to formulate field-equations for neo-Newtonian mechanics and special-relativity has been questioned by Stachel [Stachel, 1993]. This is another discussion however.

sive transformations, while within solution point transformations are active transformations.

It is not uncommon to read that in general relativity, the introduction of gravity is responsible for "restoring" symmetries that were not present in neo-Newtonian mechanics or special relativity. For instance, 't Hooft claims that "local symmetry can be restored only by adding a new field to the theory; in general relativity the field is of course that of gravitation" ['t Hooft, 1980, p. 103]. Similarly, Kosso claims that "we can add to the physics a claim about a specific force that restores the invariance" [Kosso, 2000, p. 90]. If one does not distinguish clearly between active and passive transformations, and if one is unaware that the Christoffel symbols play an important dual role in GR (as explained in chapter 4), then these statements can appear very confusing. As we have explained, by employing a generally covariant formalism, GR does (in a weak sense) extend the symmetries of previous spacetime theories which are not formulated in a generally covariant way. A result of this shift to the generally covariant formulation is the appearance of the Christoffel symbols in the equations of motion. These Christoffel symbols are also used in GR to represent the gravitational field. However, one should be careful to separate their role as implementing general covariance from their role in representing the gravitational field. In other words, GR is not a theory of gravitation because it is generally covariant.[9] As generally covariant formulations of neo-Newtonian mechanics and special relativity show, it is possible to extend the passive symmetries of a theory without creating a new theory that represents a dynamical force of nature (like gravitation). Furthermore, when it is claimed that gravity "restores" local symmetry, the relevant restoration applies only to passive transformations. While general covariance ensures that all coordinate transformations are symmetries in a weak sense, it certainly does not entail that all point transformations are symmetries. The fact that GR is a theory of gravitation also has no generalizable consequences for the class of point transformations that are symmetries. In GR, the active symmetry transformations depend on the particular metric at hand, and in general there will be no active symmetry transformations. The following table summarizes the results of this section.

[9] I am tempted to add the converse, namely that GR is not generally covariant because it is a theory of gravitation. However, this statement may be incorrect and depends on whether or not it is possible to formulate a non-generally covariant theory which makes the same predictions as GR. Newtonian mechanics is a non-generally covariant theory of gravitation, but it famously does not make the same predictions as GR.

	Active/Passive	Global/Local	Symmetry
Coordinate transformations (all)	Passive	Both	Weak: Yes; Strong: depends on $g_{\mu\nu}$
Point transformations (with dragged-along tensor fields)	Passive	Both	Yes (of EFE's)
Point transformations (within-solution)	Active	Both	Depends on $g_{\mu\nu}$

5.2.2 Quantum mechanics

As we showed in chapter 3, transformations in quantum mechanics take the form of maps between Hilbert spaces, or maps of Hilbert spaces onto themselves. Each observer has an associated Hilbert space. Two different observers O and $\bar{O}$ will each describe the state of a system S by vectors in the identical Hilbert spaces $\mathscr{H}_1$ and $\mathscr{H}_2$ respectively. A translation from the description of O to the description of $\bar{O}$ is a passive transformation, and is represented by a map $T : \mathscr{H}_1 \to \mathscr{H}_2$. Since $\mathscr{H}_1$ and $\mathscr{H}_2$ are copies of the same Hilbert space, the passive transformation from the description of O to the description of $\bar{O}$ is often just represented as a map from $\mathscr{H}_1$ onto itself and can therefore be represented by an operator on the space. An active transformation changes the state of a system as described by a single observer O. It is therefore also represented by a map of the Hilbert space onto itself.

In spacetime theories, the structure of spacetime itself determines which passive transformations are strong symmetries, and therefore which active transformations are symmetries. In addition to these, there are also weak passive symmetries which depend on the precise formulation of the theory that is adopted. In GR, the structure of spacetime is determined by the EFE's. In special relativity, the structure of spacetime is determined by the Relativity Principle and the Light Principle. Similarly, it is possible to impose certain restrictions on the quantum mechanical formalism such that it respects the symmetries that we believe to hold in the world. Remarkably, imposing such restrictions allows successful predictions to be derived from the quantum formalism, for instance about atomic structure, selection rules for electron transitions and conservation laws. Wigner is famous for his work in this domain [Gross, 1995]. Non-relativistic quantum mechanics takes place on flat spacetime, and therefore the restrictions that are imposed on the formalism ensure that transformations representing elements of the Poincaré group are symmetries. These restrictions make the observations of appropriately related observers equivalent and can therefore be classified as strong symmetries. Consistent with our analysis, these restrictions can be formulated equivalently as active or passive versions. There is no agreement on how to formulate a generally covariant version of quantum mechanics, and therefore the notion of a weak symmetry does not apply here.

I will now give an example of the restrictions that can be imposed on the quantum formalism to make it exhibit certain symmetries. I will start with the passive case. Consider that O and $\bar{O}$ are two observers related by a transformation corresponding to an element of the Poincaré group. They both describe the evolution of a system S using the Schrödinger equation. Thus, for O the evolution of the system will be

$$i\hbar\frac{d}{dt}|\phi_O\rangle = \hat{H}(t)|\phi_O\rangle \tag{5.7}$$

and for $\bar{O}$ it will be

$$i\hbar\frac{d}{dt}|\phi_{\bar{O}}\rangle = \hat{\bar{H}}(t)|\phi_{\bar{O}}\rangle \tag{5.8}$$

In order to respect the fact that the transformation from O to $\bar{O}$ is a symmetry, we impose the restriction that $\hat{\bar{H}} = \hat{H}$ [Fonda and Ghirardi, 1970, p. 30]. This implies that the time-evolution operators for both observers must be equal $\hat{\bar{T}}(t, t_0) = \hat{T}(t, t_0)$.

A similar example can be given in the case of an active transformation. Consider an observer O describing a system S and an observer $\bar{O}$ describing a system $\bar{S}$, where $\bar{O}$ is related to $\bar{S}$ in the same way that O is related to S. Being related "in the same way" means that "the expectation values of the operators of the considered irreducible set coincide for O and $\bar{O}$ respectively" [Fonda and Ghirardi, 1970, p. 36]. This means that when O makes a measurement on S at t_0, the probability that she finds a certain outcome is the same as when $\bar{O}$ makes the same measurement (represented by the *same* operator, see section 3.2) on $\bar{S}$ at her time t_0. S and S' are exact copies of each other (for instance they may be prepared using the same apparatus), and therefore can be treated as the same system in two different states.

In the time interval $[t, t_0]$, S evolves and O can experimentally determine transitions probabilities for the system, given by

$$P_{fi}(t, t_0) = |\langle\phi_f|\hat{T}(t, t_0)|\phi_i\rangle|^2 \tag{5.9}$$

Similarly, $\bar{O}$ can determine the transition probabilities for $\bar{S}$, which are given by

$$\bar{P}_{fi}(t,t_0) = |\langle\phi_f|\hat{\bar{T}}(t,t_0)|\phi_i\rangle|^2 \tag{5.10}$$

If S and $\bar{S}$ are related by a transformation corresponding to an element of the Poincaré group, then one can impose the restriction that $P_{fi}(t,t_0) = \bar{P}_{fi}(t,t_0)$. This entails that $\hat{T}(t,t_0)$ is equal to $\hat{\bar{T}}(t,t_0)$ up to a phase factor, which can be eliminated by further considerations [Fonda and Ghirardi, 1970, p. 40].

Due to the way that transformations are defined in quantum mechanics, the issue of the transformation condition does not arise. Active transformations are explicitly stated to relate different physical systems. The following table summarizes the results of this section.

	Active/Passive	Global/Local	Symmetry
Maps of $\mathscr{H}$ onto itself corresponding to elements of the Poincaré group	Active	Global	Yes
Maps of $\mathscr{H}_1$ onto $\mathscr{H}_2$ corresponding to elements of the Poincaré group (with $\mathscr{H}_1$ and $\mathscr{H}_2$ copies of $\mathscr{H}$)	Passive	Global	Yes (strong)

5.2.3 Gauge theories

As we saw in chapter 2, Wigner believed that gauge transformations could only be interpreted passively, and that they were thus very different from the spacetime transformations of the Poincaré group. This view is still dominant in the literature today, as can be seen in [Brading and Brown, 2004] and [Healey, 2009]. Wallace and Greaves have recently suggested a different view, which will agree with my analysis in this section in many respects, although some differences will remain [Wallace and Greaves, 2011]. In what follows I will classify the symmetries of gauge theory first in the context of their action on classical systems and then in the context of their action on quantum systems. I will show that active gauge transformations are possible. Where appropriate, I will show how the transformations can be represented geometrically in the fibre bundle formalism. This will be helpful in distinguishing the active and passive interpretations of gauge transformations, just as the distinction between coordinate transformations and point transformations was helpful in the context of spacetime theories.

Gauge transformations on classical systems

In classical electromagnetism, a gauge transformation is a transformation of the electric and magnetic potentials that takes the form

$$\begin{cases} \phi \to \phi' = \phi - \frac{\partial \chi(\mathbf{x},t)}{\partial t} \\ \mathbf{A} \to \mathbf{A}' = \mathbf{A} + \nabla \chi(\mathbf{x}, t) \end{cases} \tag{5.11}$$

where $\chi(\mathbf{x}, t)$ is a smooth function of space and time. As shown in chapter 2, this transformation is a symmetry because the electric and magnetic fields derived from the transformed potentials are the same as those derived from the untransformed potentials. We know that it can be interpreted passively, but in order to make a clear analogy with the spacetime case, it would be advantageous to have some interpretation of the transformation as the change of something like a coordinate system. The fibre bundle formulation of classical electromagnetism allows us to do this.

A fibre bundle formulation of classical electromagnetism Using the mathematical notions introduced in section 3.3, electromagnetism can be given a geometrical formulation. Classical Maxwell theory can be represented by the geometry of a principal fibre bundle $(E, \pi, \mathcal{M})$ with structure group $U(1)$ [Nakahara, 2003, p. 399]. $U(1)$ is an Abelian, one-dimensional Lie group, which means that it is a one-dimensional manifold and that the elements of the group commute. The base space $\mathcal{M}$ is then taken to represent space, or spacetime. If $\mathcal{M}$ is topologically equivalent to $\mathbb{R}^3$ or $\mathbb{R}^4$, then the bundle is trivial, and by theorem (2) it has a global continuous cross-section. The connection on the fibre bundle ω is the mathematical object that encodes the presence or absence of electromagnetic effects, in other words, the states of the $\mathbf{E}$ and $\mathbf{B}$ fields.

The connection ω can be related mathematically to the electromagnetic potential A_μ. In general, this can only be done locally in a open subset U of $\mathcal{M}$. However, because in this case the bundle is trivial, it is possible to define a global relation between ω and A_μ. In order to achieve this, it is necessary first to define a cross-section $\sigma : \mathcal{M} \to E$ of the bundle. This process corresponds to arbitrarily choosing a particular point in the fibre above each point of the base space. Once a section has been chosen, it is possible to define the vector potential as (Isham [1999], p.259)

$$iA_\mu(x) = (\sigma^* w)_x(\partial_\mu) \tag{5.12}$$

The operation σ^* is the pull-back, and the ∂_μ represents the local vector field. Although this equation contains some undefined notions, it is important to us because it shows that the electromagnetic potential can be encoded geometrically in terms of a connection on a principal fibre bundle. Relevant to our search for an analogue of the coordinate system is the fact that the geometric structure of the principal bundle doesn't uniquely determine the electromagnetic potential. In order to extract a precise value for A_μ from the connection ω one must make a choice of section σ. If a different section σ' had been chosen, this would have produced a different electromagnetic potential A'_μ. It can be shown that under a change of section, A'_μ and A_μ are related by a transformation which reduces to the standard gauge transformation [Isham, 1999, p. 259]

$$A'_\mu = A_\mu + \partial_\mu \chi(x) \tag{5.13}$$

where $\chi(x)$ is some smooth function of space or space and time depending on the nature of the base space. The **E** and **B** fields can then be calculated in the usual way with A_μ. The worldlines of particles in spacetime are determined by Lorentz's force law.

The fibre bundle structure of the electromagnetic field suggests that passive gauge transformations should be interpreted as changes in the section of the principal bundle. This is similar to the way that passive transformations in spacetime theories are changes in the coordinatization of the manifold. Two characteristic features of coordinate systems are their necessity and their arbitrariness. They are necessary in order to have a description of the events unfolding in the manifold that can be related to the measurements of a physical observer. On the other hand, they are arbitrary because there are many different possible coordinatizations of the manifold. These two characteristics are also features of sections on the principal bundle. A section is necessary in order to obtain a value for the electromagnetic potential that can be used by an observer making measurements. On the other hand, many different choices of this section are possible. The analogy between cross-sections and coordinate systems can also be given a mathematical justification. If a spacetime theory is formulated as a $GL(m, \mathbb{R})$-bundle $\mathbf{B}(\mathcal{M})$ of frames on an m-dimensional manifold $\mathcal{M}$, then a local choice of section on $U \subset \mathcal{M}$ given by $\sigma : U \to \mathbf{B}(\mathcal{M})$ corresponds to the choice of a local coordinate system (U, ϕ) on $\mathcal{M}$ [Isham, 1999, p. 261].

The analogy between coordinate systems and sections of the principal bundle is not unanimously accepted. Leeds, who defends a view similar to ours in the context of Q-EM, thinks nevertheless that the analogy cannot apply in classical EM. He writes [Leeds, 1999, p. 607]

> Now, I am *not* about to claim that we should think of the vector potential in classical electrodynamics as a single quantity

> which the different gauges allow us to coordinatize in different ways.

His main argument is that "there is no quantity of which the various vector potentials are different coordinatizations" [Leeds, 1999, p. 610]. This is a strange claim, given that Leeds claims to take the "fibre bundle formulation of electrodynamics literally", and that we have just shown that in such a formulation, it is possible to interpret gauge-transformation related vector potentials as different "coordinatizations" of the connection ω. However, it seems that Leeds does not fully appreciate that a fibre bundle formulation of classical electrodynamics is possible. Instead, he mistakenly believes that it can only be given in the case of gauge transformations acting on quantum systems. Healey notes that the fibre bundle formulation that Leeds considers is "a little different from the usual fibre bundle formulation of classical electromagnetism" since he appeals to a "bundle of phases" which can indeed only be applied in the presence of quantum systems [Healey, 2007, p. 99]. I therefore conclude that Leeds has not fully appreciated the power of the fibre formulation to provide an analogy between changes of section on the principal bundle and coordinate transformations.

Another more potent objection is provided by Healey, who writes [Healey, 2007, p. 17]

> It is tempting to conclude that the effects of electromagnetism [...] are represented by a unique, invariant connection on the principal fibre bundle, and that a gauge transformation corresponds merely to a change from one "coordinatization" of this connection to another. Indeed this way of reading the fibre bundle formulation of electromagnetism motivates a common strategy for interpreting this and other gauge theories whose adequacy will be a major concern of this book.

However, Healey notes that gauge transformations can also be implemented by vertical bundle automorphisms, which are the gauge theory equivalent of the point transformations of spacetime theories. In order to discuss the significance of this claim, I first define the notions of a principal automorphism and a vertical automorphism.

Definition 14. [Healey, 2007, p. 239] A *principal automorphism* of a principal fibre bundle $(E, \pi, \mathcal{M})$ is a smooth map $h : E \to E$ (from the total space onto itself) satisfying $h(pg) = h(p)g$ for all $p \in E$ and $g \in G$. This means that the image of a point in the orbit of p, is mapped onto a point in the orbit of the image of p, or that h is "G-equivariant" [Isham, 1999, p. 225].

A vertical automorphism is a principal automorphism satisfying an extra condition.

Definition 15. [Healey, 2007, p. 239] A *vertical automorphism* is a principal automorphism satisfying $\pi(h(p)) = \pi(p)$. In other words, a vertical automorphism maps the fibres back onto themselves without moving them around in the base space.

Under a vertical automorphism h, the connection can also be transformed according to $\omega \to h^*(\omega) = \omega'$, where h^* is the pull-back operation and, in general, $\omega \neq \omega'$. Furthermore, given a section σ on the principal fibre bundle, the electromagnetic potentials derivable from ω and ω' are $iA_\mu(x) = (\sigma^* w)_x(\partial_\mu)$ and $iA'_\mu(x) = (\sigma^* w')_x(\partial_\mu)$, with A_μ and A'_μ related by the standard gauge transformation (5.13) [Isham, 1999, p. 260]. Thus, a vertical bundle automorphism causes a transformation in the connection $\omega \to \omega'$, but this transformation does not have physical significance because, in a given section, the electromagnetic potentials generated by ω and ω' are gauge transforms of each other, and therefore represent the same **E** and **B** fields.[10]

The important conclusion Healey draws from considering vertical automorphisms is that the same electromagnetic situation (the same configuration of **E** and **B** fields) can be given by two *different* connections ω and ω' in the principal fibre bundle. He concludes that the connection cannot be the "real" representative of electromagnetic properties, and therefore that we should not take too seriously the idea that different sections are simply different coordinatizations of the same "real electromagnetic situation" [Healey, 2007, p. 102].

Ironically, Healey's objection can be turned against him by showing that it actually strengthens the analogy with general relativity, rather than weakening it. A principal (vertical) automorphism $h : E \to E$ in gauge theory is a very similiar mathematical operation to a diffeomorphism $d : \mathcal{M} \to \mathcal{M}$ in general relativity. In both cases, the mapping can be used to transform the geometric objects (connections in gauge theory, vector or tensor fields in GR) defined in the spaces in which they act. Thus the transformed (pulled-back) connection $h^*\omega$ is analogous to the dragged-along metric and stress-energy tensors $d^* g_{ab}$ and $d^* T_{ab}$ [Isham, 1999, p.258]. In GR, the hole argument leads us to identify diffeomorphically related models, and it seems we should come to the same conclusion in gauge theory. I conclude that the diffeomorphism invariance of GR and the vertical automorphism invariance of the fibre bundle formulation of EM are analogous and thus that the analogy between sections and coordinate systems is strengthened.

[10] I think that Healey doesn't have to restrict himself to vertical automorphisms, he could also consider just principal automorphisms [Isham, 1999, p. 259].

Finally, a more general argument against the analogy between sections and coordinate systems might be that it relies on a fibre bundle formulation of classical electromagnetism and that, in the absence of quantum motivations such as the Aharonov-Bohm effect, this is simply mathematical overkill. A defender of such a view might point to the adequacy of the **E** and **B** as an ontology of classical electromagnetism, and therefore that a formulation that gives such an important role to the surplus structure A_μ is unnecessary. Such an objection is probably fueled by strong intuitions about what physical theories should look like, as well as a desire for minimal mathematical apparatus. Faced with the impossibility of conclusively refuting such intuitions, I can only remark that the search for alternative representations of our physical theories is a fruitful pursuit because it is likely to lead to new discoveries. Furthermore, it should be seen as very satisfactory that an old theory such as electromagnetism can be a given a reformulation using the same mathematical tools as the newer theories, to which they are indispensable.

Active gauge transformations The fibre bundle formalism has allowed us to clarify how gauge transformations can be interpreted as passive transformations that are analogous to the coordinate transformations of spacetime theories. I will now show that gauge transformations can also have active interpretations. Consider the gauge transformation given by the function $\chi(\mathbf{x}, t) = -kt$, with k a constant [Healey, 2009, p. 700]

$$\begin{cases} \phi \to \phi' = \phi + k \\ \mathbf{A} \to \mathbf{A}' = \mathbf{A} \end{cases} \tag{5.14}$$

This corresponds to raising the electric potential uniformly by a constant, while leaving the magnetic potential unchanged. As Healey remarks, this transformation is technically a local gauge transformation because the function $\chi(\mathbf{x}, t)$ is not a constant. However, its effect on the physics is to raise the electric potential *uniformly* at each point of space(time). Given our earlier recommendation to define the local/global distinction in terms of the effects of the transformation on the physics rather than in terms of the function defining the transformation, I conclude that this transformation would be better classified as global. Something similar is suggested by Wallace and Greaves [Wallace and Greaves, 2011, p. 20]. From our discussion of the Aharonov-Bohm effect in chapter 4, we know that the potential everywhere inside a Faraday cage (a hollow conductor) can be raised uniformly by charging the surface of the cage. It would seem therefore that the gauge transformation (5.14) can be implemented actively in a Faraday cage setup. When this transformation is interpreted passively, it is a symmetry because the

E and **B** fields remain unchanged. Nothing in the definition of these fields allows the untransformed and transformed potentials to be distinguished. We can therefore classify this transformation (and in fact all passive gauge transformations acting on classical systems) as strong symmetries. This suggests that the corresponding active transformation of Faraday's cage is also a symmetry, as indeed it is. This entails that an observer inside Faraday's cage cannot know whether the cage is charged or uncharged by making measurements confined to the interior of the cage. This corresponds to the principle of Faradean relativity discussed in chapter 1.

Healey objects to the conclusion that the Faraday cage setup can be used to implement an active gauge transformation on the grounds that, considering only classical systems, there is no physical way to distinguish the interiors of the charged and uncharged cages [Healey, 2009, p. 711]

> It is natural to describe the state of Faraday's cube when charged by saying that it has been raised to an electric potential with respect to the ground. But this is not something that we observe - all we observe are differences in electric field outside the cube when charged and uncharged.

He picturesquely dismisses the claim that the interior of the charged cage can be considered as physically different from the interior of the uncharged cage with the following claim [Healey, 2009, p. 711].

> To suppose that one can change the electromagnetic condition inside Faraday's cube by charging its exterior is just as mistaken as to think that one can move a car from New York to Los Angeles merely by selling it.

In our terminology, Healey is claiming that the transformation implemented by the Faraday cage does not pass the transformation condition. This should be contrasted with the situation inside the cabin of Galileo's ship which is changed in the case of a uniform velocity boost. Healey tries to support this conclusion by a technical argument. He asks us to consider a situation in special relativity in which there are two Galileo ships in uniform motion with respect to each other. He claims that there is no coordinate system in which both ships are at rest. This is obvious as there is a physical difference in the states of the two ships that must show up in any description of the situation.[11] However, Healey claims that things are very different in

[11] Note that in a generally covariant formulation of SR the situation is slightly more subtle. If the ships are far enough apart, it is possible to perform a local coordinate transformation such that the 3-tuples of real numbers used to describe the spatial positions of both ships are constant. However, this would not mean that the ships are at rest with respect to each other because the metric would change as a function of time. This would ensure that when the distance between the ships was calculated, it would still be found to increase.

electromagnetism. Consider two Faraday cages a distance apart from each other. One cage is charged and the other isn't. In an appendix, Healey claims to show that there is a choice of gauge such that the interiors of both Faraday cages are at the same electric and magnetic potentials [Healey, 2009, p. 718]. He concludes that there is no physical difference between the interiors of the two cages.

There are several strong reasons to be suspicious of Healey's conclusion. The upshot of his technical argument is that he has gauged away a potential difference between the interior of the charged cage and the interior of the uncharged cage. But this should be impossible, because potential differences have physical consequences that can be measured using voltmeters. I believe that the transformation Healey has derived in his appendix is not a gauge transformation because it is discontinuous at the surfaces of the Faraday cages. However, a gauge transformation must be a smooth function of space and time. Wallace and Greaves also call attention to this fact, although they do not relate their considerations to Healey's argument [Wallace and Greaves, 2011, p. 19]. I believe that Healey's arguments should be resisted, and that the Faraday cage setup can be considered to implement a global active gauge transformation.

I also believe that Faraday's cage is an electromagnetic analogue of Galileo's ship. In both cases, the physical state of an isolated subsytem of the universe is transformed uniformly with respect to the environment, and in both cases this transformation is a symmetry. I also classify both transformations as global. Wallace and Greaves agree that Faraday's cage implements an active gauge transformation, however, they disagree that it is analogous to Galileo's ship [Wallace and Greaves, 2011, p. 20]. This is because they advocate a classification of symmetries in which they pay particular attention to the way in which the boundary conditions of the isolated subsystem are affected by the transformation. The transformation in the case of Galileo's ship preserves the boundary conditions around the cabin of the ship. However, in the case of Faraday's cage, the additional charge added to the cage to implement the transformation entails that the boundary conditions are not conserved. Following their classification of symmetries, Wallace and Greaves find that Faraday's cage is in fact analogous to Einstein's lift, as described in chapter 1 [Wallace and Greaves, 2011, p. 20].

Given the classification scheme that they endorse, Wallace and Greaves' conclusion is correct. However, I believe that this classification scheme can be criticized on conceptual grounds. I do not think that Einstein's lift should be considered as a symmetry in the same way as the ship or the cage. This is because Einstein's lift is only an approximate symmetry, whereas the ship and the cage are exact symmetries. The lift is only an exact symmetry in the case of a uniform gravitational field, and in this case there is something trivial about the symmetry (all particles are given an acceleration in the same direction which is then canceled by the appropriate choice of a coordinate system).[12] As I argued in chapter 4, I believe that the real value of Einstein's lift thought experiment is as a stepping stone towards the realization that the Christoffel symbols could be used to encode the existence of a gravitational field. However, I do not believe that Einstein's lift should be taken to implement an active symmetry transformation in the same way as Galileo's ship or Faraday's cage. Instead, I believe the emphasis should be placed on the fact that in both the latter cases, the physical state of an isolated subsystem of the universe is uniformly transformed with respect to the environment. In both cases, this transformation does not affect the internal evolution of the subsystem. This is the significant way in which Galileo's ship and Faraday's cage are analogous.

Faraday's cage shows how a particular gauge transformation can be interpreted actively. The question now arises whether all gauge transformations could be given active interpretations. Whether an arbitrary gauge transformation can be interpreted actively depends on the possibility of defining an electromagnetic situation (a configuration of charges, currents, magnets etc.) such that the electric and magnetic potentials are transformed in the appropriate way. Since all passive gauge transformations are strong symmetries, we would expect any active gauge transformation to also be a symmetry. However, apart from Faraday's cage, I am aware of no other way to actively implement a gauge transformation in the classical context.[13] The results of this section are summarized in the following table.

	Active/Passive	Global/Local	Symmetry
Change of section on the principal bundle (transformation of the electric and magnetic potentials according to (5.11))	Passive	Both	Yes (strong)
Vertical bundle automorphism (with pulled-back connection)	Passive	Both	Yes
Faraday cage	Active	Global	Yes

[12]The equality of inertial and gravitational mass is clearly a non-trivial consequence of the lift thought experiment. However, this is also a very different kind of conclusion than is drawn from Galileo's ship or Faraday's cage.

[13]In a way that will shortly be made precise, I believe that the insertion of a phase shifter (or an AB coil) can be seen to implement an active gauge transformation on a quantum system. In the quantum context, these are not symmetries because local gauge transformation are demoted to weak symmetries in Q-EM. However, it is interesting to note that in the classical context, the insertion of an AB coil is a symmetry, because it does not change the values of the electric and magnetic fields outside the coil.

Gauge transformations on quantum systems

In the presence of a quantum system represented by a wavefunction $\Psi(\mathbf{x},t)$, we saw in chapter 2 that a gauge transformation takes the following form

$$\begin{cases} \psi(\mathbf{x},t) \to \psi'(\mathbf{x},t) = e^{iq\chi(\mathbf{x},t)}\psi(\mathbf{x},t) \\ \phi \to \phi' = \phi - \frac{\partial\chi(\mathbf{x},t)}{\partial t} \\ \mathbf{A} \to \mathbf{A}' = \mathbf{A} + \nabla\chi(\mathbf{x},t) \end{cases} \tag{5.15}$$

In the presence of the quantum system it becomes more delicate to determine when a transformation is a symmetry and how an active transformation can pass the transformation condition. In addition to the equations for the **E** and **B** fields, the passive gauge transformations must also leave the Schrödinger equation unchanged. In the active case, there must be no physical consequences of the transformation. As the Aharonov-Bohm effect showed, a transformation that leaves the **E** and **B** fields unchanged in the region where the wavefunction is non-vanishing can still have physical consequences if it affects the integrability of the wavefunction. Furthermore, because the physical effects of transformations on quantum systems manifest themselves as changes in interference patterns, it becomes more difficult to define active transformations as acting on isolated subsystems. The holistic or non-separable character of quantum systems poses new problems for the classification of transformations.

The non-interacting Schrödinger equation (in natural units)

$$\frac{-1}{2m}\nabla^2\psi(\mathbf{x},t) = i\frac{\partial\psi(\mathbf{x},t)}{\partial t} \tag{5.16}$$

is invariant under gauge transformations when $\chi(\mathbf{x},t)$ is a constant. In the literature, this is called a global transformation, because the phase of the wavefunction is shifted by the same amount at each point of spacetime. This can be interpreted passively as a change in the convention used to define the zero point of phase. The fact that the phase of the wavefunction at a point is unmeasurable (only phase differences can be measured) is often given as a reason for this freedom in defining the 'zero phase'. (5.16) is not invariant under gauge transformations when $\chi(\mathbf{x},t)$ is not a constant. Instead, one needs to change the Schrödinger into its 'interacting' form

$$\frac{-1}{2m}(\nabla - iq\mathbf{A})^2\psi(\mathbf{x},t) + q\phi\psi(\mathbf{x},t) = i\frac{\partial\psi(\mathbf{x},t)}{\partial t} \tag{5.17}$$

(5.17) is invariant under arbitrary gauge transformations when the electric and magnetic potentials are transformed along with the wavefunction according to (5.15). This situation should be compared to the way in which the geodesic equation of GR is "invariant" under general coordinate transformations when the Christoffel symbols are transformed along with the coordinates. In this way, the values of the electric and magnetic potentials allow us to distinguish between certain choices of gauge, which can therefore not be considered equivalent. This entails that in the case of gauge transformations acting on quantum systems, a distinction between strong passive gauge transformations and weak passive gauge transformations becomes necessary. Thus, only transformations of the form $\chi(\mathbf{x}, t) = constant$ are strong symmetries.[14] All the other gauge transformations are weak symmetries. Consistent with this analysis, we will see that only the strong symmetries imply active symmetries. The active transformations corresponding to the weak symmetries are not symmetries.

Before moving on to an analysis of active gauge transformations, I show how passive gauge transformations acting on quantum systems can be given a geometrical interpretation in terms of fibre bundles. I will show that the analogy between a section of the principal bundle and a coordinate system of a spacetime theory remains valid in this new context.

A fibre bundle formulation of quantum-electromagnetism The theory of a quantum particle interacting with a classical electromagnetic field can be given a fibre bundle formulation in terms of a vector bundle *associated* with the principal bundle. The principal bundle represents the electromagnetic field and a section of the vector bundle (a vector field) represents the phase of the quantum particle at each point of the base space (space or spacetime). The particular way in the which the two bundles are associated encodes the action of the electromagnetic field on the quantum particle. We will see that the connection on the principal bundle defines the covariant derivative of the section of the associated bundle. In this way, the equation for the vanishing of the covariant derivative of the section of the associated bundle determines how the electromagnetic situation affects the phase of the wavefunction. As will become clear, it is very important to clearly distinguish the notion of a section of the principal bundle and a section of the associated bundle. As explained in the previous section, a section of the principal bundle is analogous to a coordinate system. On the other hand, the section of the associated bundle determines the phase of the wavefunction, and is therefore physically significant.

An associated fibre bundle $(E_F, \pi_F, \mathcal{M})$ is a vector fibre bundle if the fibre F is some vector space. In the case of a quantum particle, F is a copy of the complex vector space $\mathbb{C}$. The crucial fact in relating the fibre bundle

[14] If we take the time-independent Schrödinger equation, then the possibility arises of considering $\chi(\mathbf{x}, t) = kt$ with k a constant as a strong symmetry

representing the electromagnetic field to the one representing the quantum particle is that the structure group G of the principal fibre bundle can act (via some representation) as a group of transformations on F. In the present case, G is the group $U(1)$, and thus acts from the left on $\mathbb{C}$ by performing a rotation of elements $z \in \mathbb{C}$. Both bundles have the same base space $\mathcal{M}$.

Definition 16. [Isham, 1999, p. 233] Let $\xi = (E, \pi, \mathcal{M})$ be a principal G-bundle, and let F be a left G-space. Define $E_F = E \times_G F$ where $(p, v)g := (pg, g^{-1}v)$, and define a map $\pi_F : E_F \to \mathcal{M}$ by $\pi_F([p, v]) = \pi(p)$. Then $\xi[F] = (E_F, \pi_F, \mathcal{M})$ is a fibre bundle over $\mathcal{M}$ with fibre F that is said to be associated with the principal bundle ξ via the action of the group G on F.

An important feature of this definition is the nature of the space $E_F = E \times_G F$. Points of this space are equivalence classes $[p, v]$, with $p \in E$ and $v \in F$. Two points (p_1, v_1) and (p_2, v_2) are in the same equivalence class (they are the same point of E_F) if there is some $g \in G$ such that $(p_1, v_1)g = (p_1 g, g^{-1}v_1) = (p_2, v_2)$.

The phase of a non-relativistic charged quantum particle in an electromagnetic field is represented by a cross-section $s : \mathcal{M} \to E_F$ with $s(x) = [p, v]$ [Healey, 2007, p. 15]. If the principal bundle is trivial, a global cross-section s exists. A change of section s of the associated bundle can result in a change in the physical state of the quantum system manifesting itself as a change in the interference pattern. We will see that the effect of a phase shifter can be represented in this way. The following theorem about sections of the associated bundle can help to give a better understanding of their significance by providing an alternative way of thinking about them [Trautman, 1985, p. 74].

Theorem 3. *[Isham, 1999, p. 246] If $(E_F, \pi_F, \mathcal{M})$ is an associated fibre bundle, then its cross-sections are in bijective correspondence with maps $\phi : E \to F$ that satisfy $\phi(pg) = g^{-1}\phi(p)$ for all $p \in E$ and $g \in G$. The cross-section s_ϕ corresponding to such a map ϕ is given by $s_\phi(x) = [p, \phi(p)]$ where p is a point in $\pi^{-1}(\{x\})$.*

Figure 5.1 explains how the combination of a section of the principal bundle and a section of the associated bundle allows a definite phase for the system to be assigned for every spacetime point at which the system is present. In this figure, the result of the above theorem is used to facilitate the representation of the section of the associated bundle. The main idea is that given a section s of the associated bundle, a section σ of the principal bundle picks out a unique pair $s_\sigma(x) = (p, v) = (\sigma(x), v)$ from the equivalence class $s(x) = [p, v]$. Thus, fixing both s and σ defines a map $s_\sigma : \mathcal{M} \to E \times F$, with $s_\sigma(x) = (\sigma(x), \phi_s(\sigma(x)))$.

Given a section s of the associated bundle, a change of gauge $\sigma \to \sigma'$ of the principal bundle results in the transformation $s_\sigma(x) \to s_{\sigma'}(x) =$

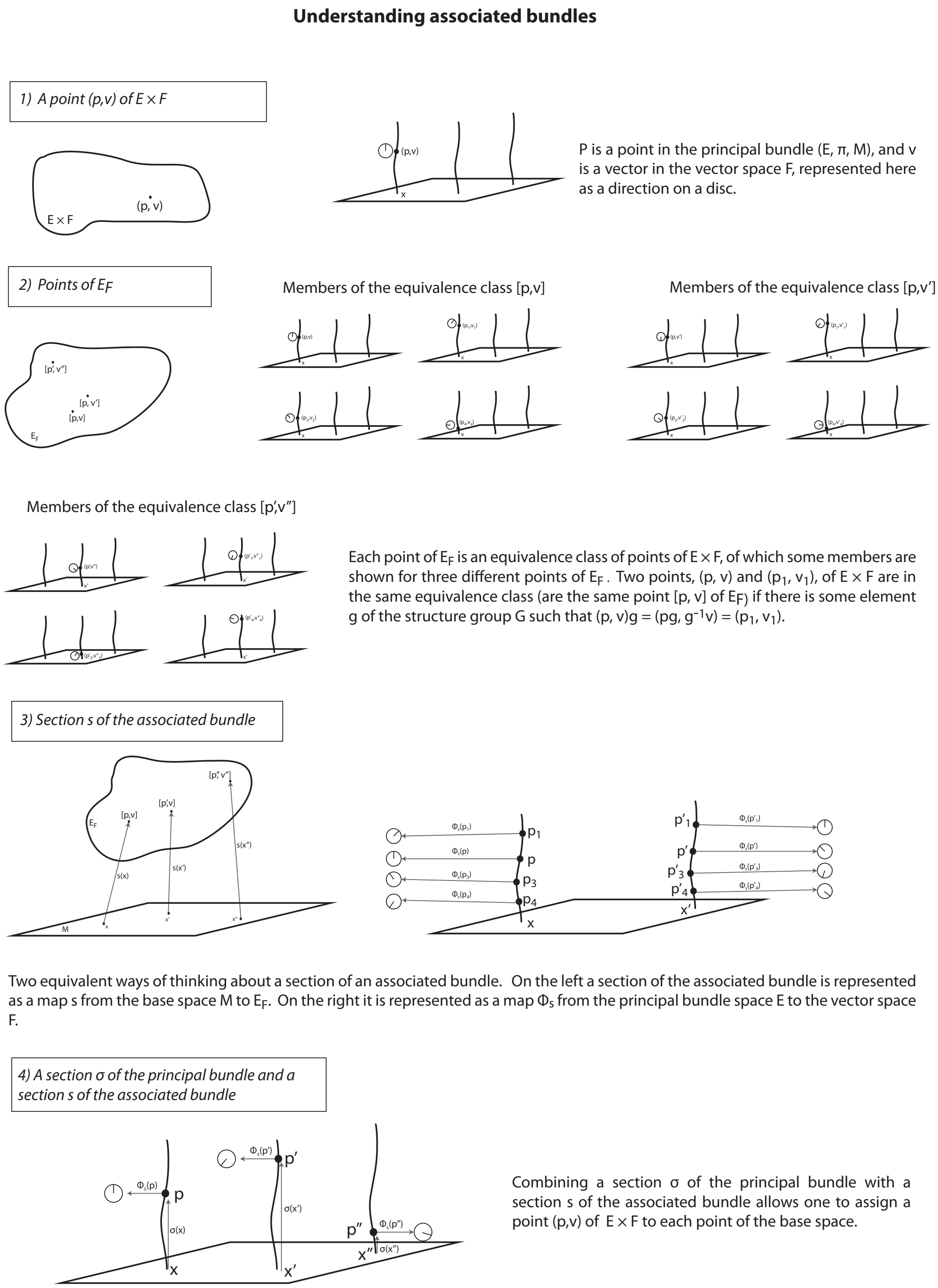

Figure 5.1: A guide to understanding associated bundles.

$(\sigma'(x), \phi_s(\sigma'(x)))$. Note that $\sigma'(x) = \sigma(x)g$, for some $g \in G$ and therefore that $s_{\sigma'}(x) = (\sigma(x)g, \phi_s(\sigma(x)g)) = (\sigma(x)g, g^{-1}\phi_s(\sigma(x))) = s(x)$, using the properties of ϕ_s. This shows that a change of section on the principal bundle preserves the equivalence classes defined by the section s on the associated bundle. Since $\phi_s(\sigma'(x)) = \phi_s(\sigma(x)g)) = g^{-1}\phi_s(\sigma(x))$, a change of section $\sigma \to \sigma'$ results in a gauge transformation of the wavefunction $\Psi(x) \to \Psi'(x) = e^{-i\theta}\Psi(x)$, with $e^{-i\theta}$ the representation of the structure group G that acts on the vector space F and θ determined by the particular group element g. Remember that a change of section $\sigma \to \sigma'$ also results in a gauge transformation of the vector potential A_μ. Thus, in the case of an associated bundle, a change of section on the principal bundle results in a joint transformation of the vector potential and the wavefunction, corresponding to (5.15). This ensures that the interpretation of a section σ on the principal bundle as analogous to a coordinate system in spacetime theories remains valid in the case of associated bundles.

Active gauge transformations acting on quantum systems Thus far, the interpretation of gauge transformations in the context of quantum systems is not significantly different from their interpretation in the context of classical systems. However, difficulties arise when trying to understand how active gauge transformations act on quantum systems. In Brown and Brading's analysis of active gauge transformations, the crux of matter seems to be whether the transformation

$$\psi(\mathbf{x}, t) \to \psi'(\mathbf{x}, t) = e^{iq\chi(\mathbf{x},t)}\psi(\mathbf{x}, t) \tag{5.18}$$

can also be interpreted actively. By considering the effects of this transformation on an electron in a double slit setup, they come to two significant conclusions.

(1) If $\chi(\mathbf{x}, t)$ is a constant, then the transformation is a symmetry, which means that it results in no change in the measurable features of the quantum system (no change in the interference pattern). Moreover, they add that the addition of a global phase to the wavefunction [Brading and Brown, 2004, p. 658]

> is of no empirical significance: physically Ψ and Ψ' represent exactly the same quantum mechanical system, indistinguishable in every way. This means that a global gauge transformation cannot be used to create an empirically distinguishable scenario.

Furthermore, they reject the possibility of applying such a global transformation to a subsystem of the universe. In order for the transformation (5.18) to have measurable effects (for it to alter an interference pattern), it must be applied to a subpart of the system itself. This is due to the particular nature of quantum systems which prohibits different systems (which are in separate quantum states) from giving rise to interference patterns. However, given that the wavefunction Ψ could represent one electron, Brading and Brown reject the idea that a part of the wavefunction could be considered as a legitimate subsystem. They conclude that the transformation (5.18) can never pass the transformation condition, and therefore that it can never have an active interpretation.

(2) They consider whether an active transformation implemented by the insertion of a phase shifter into one of the beams of the double slit experiment can be represented by the transformation (5.18), with $\chi(\mathbf{x}, t)$ not a constant over the whole of spacetime, although it may be constant in a region of spacetime. The insertion of a phase shifter results in a "relative phase transformation" given by [Brading and Brown, 2004, p. 653]

$$\psi(\mathbf{x}, t) \rightarrow \psi'(\mathbf{x}, t) = \frac{1}{\sqrt{2}}(\psi_1 e^{iq\chi} + \psi_2) \tag{5.19}$$

where ψ_1 is the beam that travels through the top slit and ψ_2 the beam that travels through the bottom slit. However, they note that a local gauge transformation of the form (5.18) can never result in a relative change in phase between the two beams, and will thus never change the interference pattern. They conclude that the insertion of a phase shifter cannot be represented by a transformation of this form, and therefore that (5.18) has no active interpretation. They give the following summary of their argument [Brading and Brown, 2004, p. 653]

> The change in the interference pattern is due to the change in the relative phase of ψ_1 with respect to ψ_2 at each point along the screen. A *local gauge transformation*, such as (5.18) will not achieve this. Local gauge freedom is the freedom to vary to overall phase of the wavefunction from point to point, but it is *not* the freedom to vary the phase of ψ_1 with respect to ψ_2 at a single space-time point.

From a purely technical point of view, there is nothing wrong with these conclusions of Brown and Brading's. However, I believe that they do not entirely do justice to the intuitive notions of what an active gauge transformation is. By very slightly adjusting the criteria for a transformation to count as an active gauge transformation, I will show that both objections of Brown and Brading's can be answered. Ultimately, I believe that the need for these adjustments is due to the particular nature of quantum systems.

I start with the first conclusion, which states that global gauge transformations can never pass the transformation condition. We have seen from our discussion of the electric Aharonov-Bohm experiment in chapter 4 that the wavefunction of a quantum particle confined to a Faraday cage undergoes a uniform phase transformation in the region of the cage when charge is added to the cage and the potential inside the cage increases uniformly. In the electric AB effect, the phase of a part of the wavefunction is transformed in this way. The subsequent change in the interference pattern is the empirical consequence of this transformation. This is evidence that a charged Faraday cage affects the phase of the wavefunction. Now consider placing the whole double slit apparatus inside a charged Faraday cage and performing the experiment. The outcome should be unchanged from when the experiment is performed in an uncharged cage. Can we not use the evidence from the electric AB effect to conclude that the double slit apparatus inside the charged Faraday cage must have undergone a global phase transformation, and therefore has passed the transformation condition? And can we not conclude that this constitutes empirical significance of global phase symmetry? Brading and Brown could point to the following passage in their paper [Brading and Brown, 2004, p. 658]

> But it is not the means by which the alleged transformation is carried out that guarantees that we have a physical transformation - it is the empirically distinct scenario.

In this way, Brading and Brown would ask for proof that when the entire double slit experiment apparatus is placed in the charged Faraday cage, the wavefunction has actually undergone a phase shift. Unfortunately, due to the nature of quantum systems, it is impossible to take a part of this wavefunction, transfer it outside of the cage and make it interfere with another wavefunction which was not inside the cage. The wavefunction inside the cage can only interfere with itself, and this is what prevents the global phase transformation performed on it from passing the transformation condition. Due to the nature of quantum systems, there are no relations between the system inside the cage and systems outside that can be used to testify that the transformation has taken place. However, we wonder in this case whether Brading and Brown are not being too rigid in their interpretation of the transformation condition. The electric AB effect provides solid empirical evidence that the phase of a wavefunction is affected by the charge on a Faraday cage. It seems unreasonable to ignore this evidence in other comparable experimental situations.

I now turn to the second conclusion. As Wallace and Greaves point out, the crux of the argument is whether an active gauge transformation must necessarily be formulated as a function "from spacetime to the gauge group" [Wallace and Greaves, 2011, p. 23]. By restricting active gauge transformations to taking the form (5.18) over the whole of spacetime, Brown and Brading implicitly assume that this must be so. Instead, I agree with Wallace and Greaves that an active gauge transformation need only take the form (5.18) in a *region* of spacetime. As we will shortly show, when such a transformation takes place, its effects propagate through spacetime, with the consequence that the initial and final states of the system *cannot* be related by a transformation of the form (5.18). Once again, this is a consequence of the nature of quantum systems. As Wallace and Greaves note [Wallace and Greaves, 2011, p. 23]

> [...] what is given, when we are given the pre- and post-transformed states of the universe, is *not* a function from spacetime to the gauge group, but merely the effect of whatever transformation is being performed on the particular pre-transformation (universe) state.

The fibre bundle formulation is particularly helpful in illustrating how a phase transformation in a region of spacetime will have effects that propagate out and ultimately change the interference pattern.

A local phase transformation in the fibre bundle formalism In our presentation of the fibre bundle formulation of quantum-electromagnetism, we showed that the phase of the wavefunction was represented by a section s of the associated bundle. However, we did not explain how s is determined by the electromagnetic situation encoded in the connection ω on the principal bundle. This is achieved by defining a notion of parallel transport for s which depends on ω. We then demand that s be chosen in such a way that it is parallel transported along all curves in the base space. In chapter 3, we showed that the connection ω could be used to define the horizontal lift of a curve in the base space. The horizontal lift is an intrinsic property of a curve in the principal bundle, which means it does not depend on a choice of section σ on the principal bundle. This makes it a suitable notion for defining the parallel transport of the section s. A section s is said to parallel transport the phase along a curve in the base space if it is constant along the horizontal lift of this curve [Nakahara, 2003, p. 391]. This is formalized as follows. Consider a curve $\alpha(t) : [a, b] \to \mathcal{M}$ in the base space. The horizontal lift of this curve defined by the connection ω is the curve $\alpha^{\uparrow}(t) : [a, b] \to E$. Since a section s can be represented as a map $\phi_s : E \to F$, we say that s parallel transports the phase if $\phi_s(\alpha^{\uparrow}(t))$ is constant for $t \in [a, b]$. Combining this notion of parallel transport with a section σ of the principal bundle, it is possible to assign a phase to each point of the curve $\alpha(t)$, as is shown in 5.2.

A local phase transformation in a region of space(time) can be represented by transformation of the section $s \to s'$ in this region. Let us say that a phase shifter in a region U of the base space causes a deviation from parallel

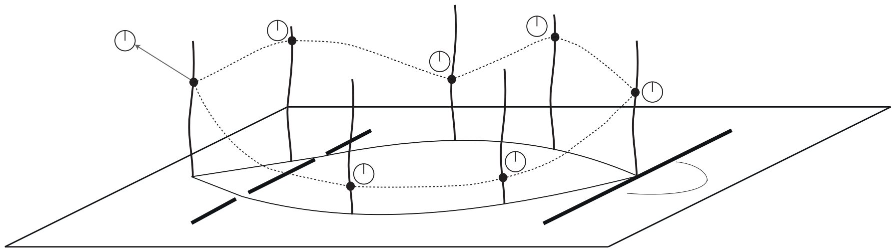

Figure 5.2: A figure showing how the phase (given by the section s on the associated bundle) is parallel transported along the horizontal lift, represented by the dotted line. Note that the dotted line closes here, and thus that the principal bundle has trivial holonomies.

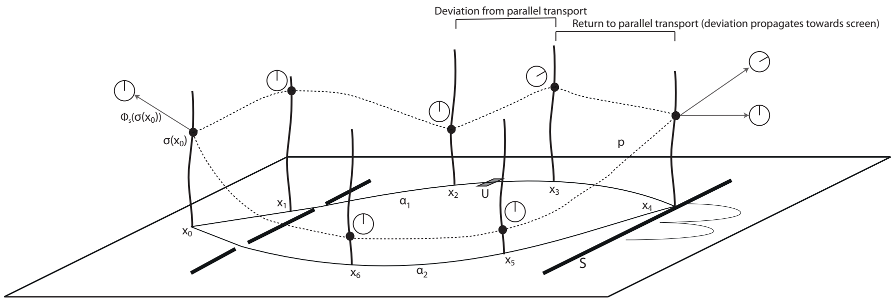

Figure 5.3: A figure showing how a phase shift in the region U, represented by the gray square, is propagated forward towards the screen. The dotted line represents the horizontal lift of the curve in the base space. Note how the value of the phase at the point x_4 on the screen has become path dependent. In fact, the effect of the phase shifter is to make the phase non-integrable, as was discussed in the case of the AB effect in chapter 4. In order to ensure that the wavefunction stays single-valued, the phases of the two different paths must interfere and a change in the interference pattern results.

transport in this region. Consider how the phase of the wavefunction evolves along a curve that passes through the region U. Let $\alpha(t)$ enter the region at $t = t_1$ and exit the region at $t = t_2$. At the entrance to the region, the phase is given by $\phi_s(\alpha^{\uparrow}(t_1))$. After passing though the region, the phase is given by $\phi_{s'}(\alpha^{\uparrow}(t_2))$. Since the phase is not parallel transported through this region, $\phi_s(\alpha^{\uparrow}(t_1)) \neq \phi_{s'}(\alpha^{\uparrow}(t_2))$. However, from t_2 to b, the phase is once again parallel transported, which means that, when it gets to the screen, the phase must be the same as when it exited the region of the phase shifter, in other words $\phi_{s'}(\alpha^{\uparrow}(b)) = \phi_{s'}(\alpha^{\uparrow}(t_2))$. This shows that the effect of the phase shifter is propagated forward toward the screen. This is illustrated in 5.3.

The effect of the phase shifter is to make the phase of the wavefunction path-dependent. Along curves that travel from the source to the screen through the region U, the phase will be shifted. However, along curves with the same end-points that do not travel through U the phase will not be shifted. In order to avoid concluding that the wavefunction must be multiply-valued, one concludes, as in the case of the AB effect, that the phases of these different paths must interfere. In the region U, the change in the phase can be represented by the transformation $\psi(\mathbf{x},t) \to \psi'(\mathbf{x},t) = e^{-iq\chi(\mathbf{x},t)}\psi(\mathbf{x},t)$, with $\chi(\mathbf{x},t)$ essentially uniform across U, dropping rapidly but smoothly to zero on the boundaries of U. As we have shown, the effect of this transformation will propagate towards the screen and will result in a relative phase shift between beams that travel through the region of the phase shifter and those that don't. From this we conclude that local active gauge transformations are possible, and will have empirical consequences. They are therefore not symmetries.

The classification of gauge transformations acting on a quantum system are given in the following table.

	Active/Passive	Global/Local	Symmetry
Global phase shift as consequence of change of section σ on the principal bundle	Passive	Global	Yes (strong)
Other phase shifts as result of change of section on the principal bundle	Passive	Local	Yes (weak)
Phase shift caused by Faraday cage	Active	Global	Yes
Local phase shift caused by change of section on the associated bundle	Active	Local	No

5.3 Conclusion: On classical and quantum systems

In this chapter I have proposed a classification of transformations that applies in three different theoretical contexts: classical spacetime theories, quantum mechanics and gauge theory. I have shown that in all three contexts transformations have a dual interpretation in terms of active and passive versions. Where possible, I have illustrated this duality with the help of a geometrical formulation of the theory. I conclude, against Wigner and Brown and Brading, that the global spacetime symmetries of the Poincaré group are not unique in having active interpretations.

Nevertheless, I concede that the arguments of Brown and Brading, even though they do not lead ultimately to the correct conclusions, highlight some important differences in the nature of classical and quantum systems. I believe that it is because of these differences that it has proven more difficult to identify how gauge transformations on the latter can be active. Firstly, the fact that quantum systems will only exhibit interference effects with themselves prevents active global transformations on them from having empirical consequences in their relations with the environment. This is not the case for classical systems. Secondly, the way in which local transformations of the wavefunction in situations such as the double slit experiment can propagate out to affect later states of the system means that the pre- and post-transformed states of the system cannot necessarily be related by a smooth function on spacetime. Once again, this is not the case for classical systems. However, we should not let these differences distract us from the fact that, in both cases of global and local gauge transformations, one can make sense of the notion of an active gauge transformation if one is willing to adapt one's criteria very slightly from the case of transformations acting on classical systems. Rather than taking these adjustments as a sign that active gauge transformations are fundamentally different from active spacetime transformations, we should take them to show that classical systems are fundamentally different from quantum systems.

Chapter 6

Conclusion

In chapter 1, I promised that this thesis would address two questions. I will now summarize the answers to these questions that the arguments in this thesis support.

What role do symmetry transformations play in GR and Q-EM?

In chapter 2 we saw that, in Wigner's philosophy of symmetry, active symmetry transformations act as principles that make possible the discovery of the laws of nature. By ensuring that certain features of the world will not affect the outcomes of measurements, these active symmetries allow scientists to isolate the relevant variables that will feature in these laws. In chapter 5, we showed that the logical equivalence of active and strong passive symmetry transformations implied that the latter could also share this significance. In chapter 4, we argued that weak passive transformations could play an important methodological role in enabling the construction of a geometric theory of a dynamical physical force. In both GR and Q-EM, we showed that by requiring the invariance of the equations of motion under weak passive transformations, certain mathematical objects could be introduced into the theory, namely the Christoffel symbols and the electromagnetic potential respectively. The properties of these mathematical objects could then be extended by submitting them to additional field equations. As a result these objects were able to encode the presence of dynamical forces, namely gravitation and electromagnetism. This line of reasoning, valid equally in GR and Q-EM, illuminated the initially mysterious possibility that the existence of a dynamical force could be deduced from a symmetry requirement. By arguing that the symmetry requirement is a stepping stone towards a certain mathematical representation of the dynamical force in question, the mystery disappears.

Is there a fundamental difference between spacetime transformations and *gauge transformations?*

For Wigner, the fundamental difference between spacetime transformations and gauge transformations was the possibility of interpreting the former, but not the latter, actively. This view is shared, among others, by Brown and Brading and Healey. In chapter 5, I showed that active gauge transformations are possible if one is prepared to adjust the criteria by which one judges when an active gauge transformation has been performed. In the case of gauge transformations acting on quantum systems, I suggested that the difficulties that arise in understanding the possibility of active gauge transformations have their origins in the particular nature of quantum systems.

The answers to these two questions suggest that, with respect to the interpretation and the role of symmetry transformations in GR and Q-EM, a certain amount of unity between the two theories already exists. This is an encouraging sign that a focus on symmetry can bring us closer to the desired unification. By highlighting areas where the unity breaks down, this thesis also points towards an important obstacle to unification, namely the fundamental difference between classical and quantum systems.

Bibliography

Y. Aharonov and D. Bohm. Significance of electromagnetic potentials in the quantum theory. *Phys. Rev.*, 115:485–491, Aug 1959.

I.J.R. Aitchison and A.J.G. Hey. *Gauge Theories in Particle Physics.* Institute of Physics Publishing, second edition edition, 1989.

Jeeva Anandan. Remarks concerning the geometries of gravity and gauge fields. In *Directions in General Relativity*, volume Vol. 1, Papers in honor of Charles Misner. Cambridge University Press, 1993.

Sunny Y. Auyang. *How is quantum field theory possible?* Oxford University Press, 1995.

Gordon Belot. Understanding electromagnetism. *The British Journal for the Philosophy of Science*, 49(4):531–555, 1998.

Herbert J. Bernstein and Anthony V. Phillips. Fiber bundles and quantum theory. *Scientific American*, July 1981.

Katherine Brading and Harvey R. Brown. Are gauge symmetry transformations observable? *The British Journal for the Philosophy of Science*, 55 (4):645–665, 2004.

Katherine A. Brading and Elena Castellani. Symmetries and invariances in classical physics. In Jeremy Butterfield and John Earman, editors, *Handbook of the philosophy of Science. Philosophy of Physics.* Elsevier, 2003.

Katherine A. Brading and Elena Catellani. *Symmetries in Physics: Philosophical Reflections.* Cambridge University Press, 2003.

Harvey Brown and Katherine Brading. General covariance from the perspective of noether's theorems, August 2002. URL `http://philsci-archive.pitt.edu/821/`.

Harvey R. Brown. Aspects of objectivity in quantum mechanics. In Jeremy Butterfield and Constantine Pagonis, editors, *From Physics to Philosophy.* Cambridge University Press, 1999.

Harvey R. Brown and Oliver Pooley. Chapter 4: Minkowski space-time: A glorious non-entity. In Dennis Dieks, editor, *The Ontology of Spacetime*, volume 1 of *Philosophy and Foundations of Physics*, pages 67 – 89. Elsevier, 2006.

T.Y. Cao. Gauge theory and the geometrization of fundamental physics. In Harvey R. Brown and Ron Harré, editors, *Philosophical Foundations of Quantum Field Theory*. Clarendon Press, 1990.

R. G. Chambers. Shift of an electron interference pattern by enclosed magnetic flux. *Phys. Rev. Lett.*, 5:3–5, Jul 1960.

Bryce S. DeWitt. Quantum theory without electromagnetic potentials. *Phys. Rev.*, 125:2189–2191, Mar 1962.

Dennis Dieks. Another look at general covariance and the equivalence of reference frames. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics*, 37(1):174 – 191, 2006.

Dennis Dieks. Lecture notes on the history and foundations of spacetime theories. Universiteit Utrecht, 2011.

John Earman. Gauge matters. *Philosophy of Science*, 69(S3):pp. S209–S220, 2002.

John Earman and John Norton. What price spacetime substantivalism? the hole story. *The British Journal for the Philosophy of Science*, 38(4):pp. 515–525, 1987.

Albert Einstein. On the electrodynamics of moving bodies. In *The Principle of Relativity*. Methuen and Company, 1923.

Albert Einstein. The foundation of the general theory of relativity. In A. J. Kox, Martin J. Kein, and Robert Schulmann, editors, *The collected papers of Albert Einstein*, volume 6. Princeton University Press, 1997.

L. Fonda and G.C. Ghirardi. *Symmetry Principles in Quantum Physics*. Marcel Dekker, 1970.

Galilei Galileo. *Dialogue Concerning the Two Chief World Systems*. University of California Press, 1967.

Robert Geroch. *General relativity from A to B*. The University of Chicago Press, 1978.

David J. Griffiths. *Introduction to Electrodynamics*. Prentice Hall, 1999.

Frank Gronwald and F.W. Hehl. On the gauge aspects of gravity. E-print: arXiv:gr-qc/9602013, 1996.

David J. Gross. Symmetry in physics, wigner's legacy. *Physics Today*, 12, December 1995.

Richard Healey. *Gauging What's Real: The Conceptual Foundations of Gauge Theories*. Oxford University Press, 2007.

Richard Healey. Perfect symmetries. *The British Journal for the Philosophy of Science*, 60(4):697–720, 2009.

Chris J. Isham. *Modern Differential Geometry for Physicists*. World Scientific, 1999.

Jenann Ismael and Bas C. van Frassen. Symmetry as a guide to superfluous theoretical structure. In Katherine A. Brading and Elena Castellani, editors, *Symmetries in Physics: Philosophical Reflections*. Cambridge University Press, 2003.

D Ivanenko and G Sardanashvily. The gauge treatment of gravity. *Physics Reports*, 94(1):1 – 45, 1983.

Michel Janssen. The twins and the bucket: How einstein made gravity rather than motion relative in general relativity, May 2011. URL `http://philsci-archive.pitt.edu/8605/`.

T.W.B. Kibble. Lorentz invariance and the gravitational field. *Journal of mathematical physics*, 1961.

P Kosso. The empirical status of symmetries in physics. *The British Journal for the Philosophy of Science*, 51(1):81–98, 2000.

Stephen Leeds. Gauges: Aharonov, bohm, yang, healey. *Philosophy of Science*, 66(4):pp. 606–627, 1999.

Holger Lyre. A generalized equivalence principle. *Int. J. Mod. Phys. D*, 69, 2000. E-print: arXiv:gr-qc/0004054v2.

Klaus Mainzer. *Symmetries of nature*. Walter de Gruyter, 1988.

Christopher A. Martin. Gauge principles, gauge arguments and the logic of nature. *Philosophy of Science*, 69(S3):pp. S221–S234, 2002.

Christopher A. Martin. On continuous symmetries and the foundations of modern physics. In *Symmetries in Physics: Philosophical Reflections*. Cambridge University Press, 2003.

Charles W. Misner, Kip S. Thorne, and John Archibald Wheeler. *Gravitation*. W.H. Freeman, 1973.

Mikio Nakahara. *Geometry, Topology and Physics*. Institute of Physics Publishing, second edition edition, 2003.

John Norton. What was einstein's principle of equivalence? *Studies in History and Philosophy of Science*, 16, 1985.

John Norton. General covariance and the foundations of general relativity: eight decades of dispute. *Rep. Prog. Phys.*, 56(7), 1993.

John D. Norton. General covariance, gauge theories, and the kretchmann objection. In Katherine. A Brading and Elena Castellani, editors, *Symmetries in Physics: Philosophical Reflections*. Cambridge University Press, 2003.

Antigone M. Nounoun. A fourth way to the aharonov-bohm effect. In Katherine A. Brading and Elena Castellani, editors, *Symmetries in Physics: Philosophical Reflections*. Cambridge University Press, 2003.

Lochlainn O'Rafeartaigh. *The Dawning of Gauge Theory*. Princeton University Press, 1997.

Murray Peshkin and A Tonomura. *The Aharonov-Bohm Effect*. Springer-Verlag, 1989.

M. L. G. Redhead. Symmetry in intertheory relations. *Synthese*, 32:77–112, 1975. 10.1007/BF00485113.

Michael Redhead. The intelligibility of the universe. *Royal Institute of Philosophy Supplements*, 48:73–90, 2001.

Michael Redhead. The interpretation of gauge symmetry. In *Symmetries in Physics: Philosophical Reflections*. Cambridge University Press, 2003.

D. W. Sciama. The physical structure of general relativity. *Rev. Mod. Phys.*, 36:463–469, Jan 1964.

John Stachel. The meaning of general covariance: The hole story. In John Earman, Allen I. Janis, J. Massey, Gerald, and Nicholas Rescher, editors, *Philosophical Problems of the Internal and External Worlds: Essays on the Philosophy of Adolf Grunbaum*. University of Pittsburgh Press, 1993.

John L. Synge. *Relativity: the general theory*. North-Holland Publishing, 1960.

Gerard 't Hooft. Gauge theories of the forces between elementary particles. *Scientific American*, 242, 1980.

Paul Teller. The gauge argument. *Philosophy of Science*, 67:pp. S466–S481, 2000.

Akira Tonomura, Nobuyuki Osakabe, Tsuyoshi Matsuda, Takeshi Kawasaki, Junji Endo, Shinichiro Yano, and Hiroji Yamada. Evidence for aharonov-bohm effect with magnetic field completely shielded from electron wave. *Phys. Rev. Lett.*, 56:792–795, Feb 1986.

Roberto Torretti. *Relativity and Geometry.* Dover Publications, 1983.

Andrzej Trautman. Fiber bundles, gauge fields and gravitation. In A. Held, editor, *General Relativity and Gravitation. Vol 1. One hundred years after the birth of Albert Einstein.* Plenum Press, 1980.

Andrzej Trautman. *Differential geometry for physicists.* Humanities Pr, 1985.

Ryoyu Utiyama. Invariant theoretical interpretation of interaction. *Phys. Rev.*, 101:1597–1607, Mar 1956.

Bas C. van Fraassen. *Laws and Symmetry.* Oxford University Press, 1989.

David Wallace. Qft, antimatter, and symmetry. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics*, 40(3):209 – 222, 2009.

David Wallace and Hilary Greaves. Empirical consequences of symmetries. E-print: arXiv:1111.4309v1 [physics.hist-ph], 2011.

David Wallace and Christopher G. Timpson. Quantum mechanics on spacetime i: Spacetime state realism. *The British Journal for the Philosophy of Science*, 61(4):697–727, 2010.

Steven Weinberg. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity.* John Wiley & Sons, 1972.

Steven Weinberg. The search for unity: Notes for a history of quantum field theory. *Daedalus*, 106(4):pp. 17–35, 1977.

Steven Weinstein. Gravity and gauge theory. *Philosophy of Science*, 66:pp. S146–S155, 1999.

Eugene Wigner. The unreasonable effectiveness of mathematics in the physical sciences. *Communications in Pure and Applied Mathematics*, 13(1), February 1960.

Eugene Wigner. *Symmetries and reflections.* The MIT Press, 1970.

Eugene Wigner. The role and value of symmetry principles and einstein's contribution to their recognition. In Colin D. Froggatt and Holger B. Nielsen, editors, *Origin of Symmetries.* World Scientific, 1991.

Eugene Wigner. The meaning of symmetry. In Jadgish Mehra, Arthur S. Wightman, and Gérard G. Emch, editors, *The collected works of Eugene Paul Wigner*. Springer, 1992a.

Eugene P. Wigner. On kinematic and dynamic laws of symmetry. In *The Collected Works of Eugene Paul Wigner*. Springer, 1992b.

Tai Tsun Wu and Chen Ning Yang. Concept of nonintegrable phase factors and global formulation of gauge fields. *Phys. Rev. D*, 12:3845–3857, Dec 1975.

A. Zee. *Quantum Field Theory in a Nutshell*. Princeton University Press, second edition edition, 2010.