Stringed along or caught in a loop?
Stringed along or caught in a loop?

*Philosophical reflections on modern quantum gravity research*

Keizo Matsubara
Abstract

A number of philosophical questions, all connected to modern research in quantum gravity, are discussed in this dissertation.

The goal of research in quantum gravity is to find a quantum theory for gravitation; the other fundamental forces are already understood in terms of quantum physics. Quantum gravity is studied within a number of different research programmes. The most popular are string theory and loop quantum gravity; besides these a number of other approaches are pursued.

Due to the lack of empirical support, it is relevant to assess the scientific status of this research. This is done from four different points of view, namely the ones held by: logical positivists, Popper, Kuhn and Lakatos. It is then argued that research in quantum gravity may be considered scientific, conditional on scientists being open with the tentative and speculative nature of their pursuits. Given the lack of empirical progress, in all approaches to quantum gravity, a pluralistic strategy is advised.

In string theory there are different theoretical formulations, or dualities, which are physically equivalent. This is relevant for the problem of underdetermination of theories by data, and the debate on scientific realism. Different views on the dualities are possible. It is argued that a more empiricist view on the semantics of theories, than what has been popular lately, ought to be adopted. This is of importance for our understanding of what the theories tell us about space and time.

In physics and philosophy, the idea that there are worlds or universes other than our own, has appeared in different contexts. It is discussed how we should understand these different suggestions; how they are similar and how they are different. A discussion on, how and when theoretical multiverse scenarios can be empirically testable, is also given.

The reliability of thought experiments in physics in general and in quantum gravity in particular is evaluated. Thought experiments can be important for heuristic purposes, but in the case of quantum gravity, conclusions based on thought experiments are not very reliable.

Keywords: quantum gravity, string theory, loop quantum gravity, philosophy of physics, philosophy of science, scientific realism, structural realism, multiverse, thought experiments

Keizo Matsubara, Uppsala University, Department of Philosophy, Logic and Metaphysics, Box 627, SE-751 26 Uppsala, Sweden.

© Keizo Matsubara 2013

urn:nbn:se:uu:diva-185554 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-185554)

Printed by Elanders Sverige AB, 2012
Dedicated to my parents,
Akira and Carine
First of all I would like to thank Lars-Göran Johansson, my advisor and collaborator. He has been very helpful and inspiring during the years we have been working together. Riksbankens Jubileumsfond funded the project, in which we started our joint work on philosophical issues regarding string theory. I am very grateful for the financial support they provided. In this project, I was hired as a researcher in my role as a theoretical physicist, since I already had a Ph.D. in theoretical physics. Shortly after being hired to work on this project, I finished my MA in philosophy. It was then decided that I was also going to be hired as a Ph.D.-student in philosophy, with the intention of using my work in the project as part of my dissertation. Lars-Göran supported this decision and for this I am very happy.

When the project, financed by Riksbankens Jubileumsfond, was finished I went to the University of Toronto. This was made possible due to a generous stipend which was given to me by The Sweden-America Foundation and donated by Borgrättsfonderna. I had a great time in Toronto, much due to my wonderful host Jim Brown. He introduced me to the interesting topic of studying thought experiments in science and philosophy for which I will be forever grateful.

I would also like to acknowledge a number of people who have contributed to my research in various ways. They have helped me through valuable discussions, by giving comments to manuscripts, by writing letters of recommendation etc. These people are: Ulf Danielsson, Richard Dawid, Sven Ove Hansson, Sten Lindström, George Masterton, Chris Smeenk, Kim Solin, Mårten Stenmark and Harald Wiltsche.

I would also like to thank all other friends and colleagues, in Uppsala, Toronto and elsewhere, who in a more indirect way have contributed to my research. They may or may not have directly influenced the contents of this thesis. Instead they have all been very important in other ways namely by making my days more enjoyable, or by helping me in various practical matters. It is very many who fall into this category and I hope they will not be sad by the fact that I have decided not to include a list with all their names. With regards to this large group of people I use the old cliché; no one mentioned, no one forgotten.

Finally I extend my deepest gratitude to my family and in particular my parents, Akira and Carine, for their constant support during all these years. I have dedicated this dissertation to them.
## Contents

Acknowledgements ........................................................................................................... vii

1 Introduction ....................................................................................................................... 13

2 Physics and Philosophy ................................................................................................... 18
   2.1 What physicists say about philosophy ................................................................. 18
   2.2 Why philosophy is relevant for physics ............................................................... 20

3 Approaches to quantum gravity ....................................................................................... 22
   3.1 Why do we need a theory of quantum gravity? .................................................. 22
   3.2 String theory ............................................................................................................ 24
   3.3 Loop quantum gravity ............................................................................................. 27
   3.4 Other approaches to quantum gravity ...................................................................... 28

4 Methodology and research in quantum gravity ............................................................... 31
   4.1 String theory ............................................................................................................ 31
   4.2 String theory and philosophy of science ............................................................. 33
      4.2.1 String theory from a logical positivist’s perspective .................................. 33
      4.2.2 String theory from a Popperian perspective ................................................. 34
      4.2.3 String theory from a Kuhnian perspective ................................................... 37
      4.2.4 String theory from a Lakatosian perspective .............................................. 39
   4.3 Further remarks concerning string theory and scientific norms ................. 43
      4.3.1 Why are tests of string theory required? ..................................................... 45
      4.3.2 String theory from an internalist perspective: explanations ...................... 46
      4.3.3 String theory from an externalist perspective ............................................. 50
   4.4 Other approaches to quantum gravity ................................................................. 51
   4.5 Conclusions .............................................................................................................. 51

5 Realism, semantics and scientific theories ........................................................................ 56
   5.1 Introduction .............................................................................................................. 56
   5.2 Earlier views on scientific theories ......................................................................... 57
      5.2.1 Logical positivism and verificationist semantics ..................................... 57
      5.2.2 Quine and semantic holism ......................................................................... 58
      5.2.3 Arguments for scientific realism ................................................................. 59
      5.2.4 Arguments for scientific anti-realism ......................................................... 60
   5.3 A literal understanding of theories? ........................................................................ 62
   5.4 Final comments ....................................................................................................... 63
## 6 Realism, underdetermination and string theory dualities

### 6.1 Introduction

### 6.2 Underdetermination and scientific realism

#### 6.2.1 Underdetermination

#### 6.2.2 Scientific theories and realism

### 6.3 Formalisms and physical content

#### 6.3.1 On how the word ‘duality’ will be used

#### 6.3.2 Purported dualities outside of string theory

### 6.4 Dualities in string theory

#### 6.4.1 T-duality and mirror symmetry

#### 6.4.2 S-duality

#### 6.4.3 AdS/CFT

#### 6.4.4 Some general remarks on the dualities

### 6.5 Different views on string theory dualities

### 6.6 Conclusions

## 7 Quantum gravity semantics and the nature of spacetime

### 7.1 Introduction

#### 7.1.1 From classical to quantum

#### 7.1.2 Background-independence

### 7.2 Physical content and physical concepts

### 7.3 Summary and conclusions

## 8 The many worlds of physics and philosophy

### 8.1 Introduction

### 8.2 A few comments on terminology

### 8.3 Possible worlds in philosophy

### 8.4 Frameworks, theories and models

### 8.5 String theory and the Landscape

### 8.6 The many worlds interpretation of quantum mechanics

### 8.7 Tegmark’s taxonomy of multiverses

### 8.8 The anthropic principle

### 8.9 Summary and conclusions

## 9 Thought experiments and quantum gravity

### 9.1 Introduction

### 9.2 The anatomy of a thought experiment

### 9.3 A taxonomy of thought experiments

#### 9.3.1 Destructive thought experiments

#### 9.3.2 Constructive thought experiments

#### 9.3.3 Platonic thought experiments

### 9.4 Brown’s platonism

### 9.5 Norton’s view that thought experiments are arguments

### 9.6 A critique of Norton’s view

### 9.7 An alternative to platonism
1. Introduction

Established contemporary physics acknowledges the existence of four forces or interactions, which are considered to be fundamental. These are: gravity, electromagnetism, the strong nuclear interaction and the weak nuclear interaction. Among these only gravity has not been satisfactorily described in terms of a quantum theory. The others are treated in the standard model of particle physics, with the use of the conceptual framework of quantum field theory, or QFT. Gravity on the other hand is described in terms of Einstein’s general theory of relativity, henceforth abbreviated GTR, which is not a quantum theory and uses a completely different conceptual framework. The use of two completely different theoretical frameworks, as the basis for physics, seems quite unsatisfactory to many physicists; they would much prefer a more unified picture. Furthermore the theories do not work together. For instance, we can not predict what would happen in many situations when we need to take both gravitational and quantum factors into account. Examples where we need a better theoretical understanding are: at the very early stages of the universe and the final stages of the life of a black hole.

The standard view on this issue is that we need a quantum theory for gravity. Einstein’s GTR would then be understood as an approximate classical version of this quantum theory in the same way as classical electrodynamics can be seen as an approximation to quantum electro dynamics, or QED.

A number of different approaches, or research programmes, have been proposed to solve the problem of developing a satisfying theory of quantum gravity. The most popular approach among present day physicists is string theory. String theory is also the main focus of this dissertation, however other approaches to quantum gravity will also be discussed. The largest among these alternative approaches is loop quantum gravity, abbreviated LQG.

The title I have chosen for this dissertation, Stringed along or caught in a loop? – Philosophical reflections on modern quantum gravity research, might suggest that my attitude to research in quantum gravity is primarily negative. This is however not the case. I would like to emphasize from the start that, while I am critical towards some aspects of the present day situation in this field of research, I am not at all against research in quantum gravity, nor am

---

1While the spelling used in this dissertation is American English, I have decided to spell “research programme” in the British fashion. This is because it is used in the specific sense described in (Lakatos 1970) later on in the text. Here this does not really matter and a looser reading of the expression “research programme” can be used.
I strongly against any particular approach to the problem of formulating a working theory of quantum gravity.

The main problem, with all approaches to quantum gravity, is that at present none of them are given any support by new experimental findings. No new precise predictions of any of the theories have been experimentally confirmed. Some experiments have been conducted, but they have all been negative. Also those experiments have been of such a nature that the theory only predicts what might happen according to some specific scenario, which is not assumed to be the only possible according to the theory. Hence the experiments are not of the form, where a theory makes a strict testable prediction and where failure to confirm the prediction give us strong grounds for abandoning the theory.

Many have criticized the various approaches to quantum gravity. String theory, being the most popular approach, has probably received most of the criticism. The lack of progress on the empirical level is the main reason for why string theory is being criticized. It is nonetheless important to point out that no other research programme has made any empirical progress either.

Among the approaches to quantum gravity there is one thing, that is often mentioned, that sets string theory apart. As opposed to the other approaches, there is a genuine unification of the different forces, not just a quantum description of gravity. So string theory is a more ambitious project. There are different opinions as to whether that is a good or a bad thing though.

What is the best strategy to pursue when a number of competing approaches work on the same problem; all without producing progress in terms of confirmed empirical predictions? Some critics argue that string theory is too dominant and thus is swallowing resources from other viable approaches to quantum gravity. This issue will be discussed in due course in this dissertation.

The lack of connection between the theories and any empirical results is of course disappointing. For this reason many researchers, within more traditional branches of physics, have complained. Charges have also been made that these theories are not genuinely scientific. This is a debate concerning the scientific status of the research and is an interesting issue from a philosophical point of view. It is one of the main topics of discussion in this dissertation. I will, besides this, address questions more dealing with the interpretation of the theories; once again string theory is the research programme that will be most thoroughly investigated.

This is a dissertation in philosophy of science, not physics. However my aim is to make the text relevant and readable both to physicists and philosophers. For this reason I have included some background that may be trivial to some readers. I have also avoided the use of mathematics unless it is needed for the philosophical points I want to make.
Here follows a short overview of the contents of the following chapters of this dissertation:

Chapter 2.
This chapter contains a general discussion on physics and philosophy. Many physicists have a very negative attitude towards philosophy and have argued that physics has no need for philosophy. I will argue that philosophy is of relevance to physics. I want to explain and show in which way philosophical investigations and discussions are important for physics, especially in the context of contemporary research in quantum gravity. I will also try to dispel some common misunderstandings concerning the role of philosophers of science.

Chapter 3.
Some background to the problem of quantum gravity will be given in this chapter. A number of different approaches to quantum gravity are briefly presented. String theory will be given most space but other approaches, especially loop quantum gravity, are also discussed. The background presented will be reasonably short and nontechnical; it is included in the text primarily to facilitate the discussion on the philosophical issues which are discussed in this dissertation. ²

Chapter 4.
Methodological issues concerning string theory, and other approaches to quantum gravity, are discussed in this chapter. The scientific status of string theory, and the other approaches to quantum gravity, is evaluated using a few different perspectives on scientific method. This chapter is closely based on the article, “String theory and general methodology: a mutual evaluation”, which was written in collaboration with Lars-Göran Johansson. It is published in, Studies in History and Philosophy of Modern Physics. The text in the chapter follows the article very closely, but has been rewritten so as to broaden the scope of the topic a little. Now it contains more comments on other approaches to quantum gravity, besides string theory.³

²In this chapter, parts of the articles and manuscripts on which this dissertation is largely based, have been used for the background on string theory. To avoid repetitions the material taken from the original sources has been edited and rewritten, however much of the original text remains. Of these some have been taken from texts I jointly wrote with Lars-Göran Johansson.
³Even though the text has been rewritten, it is fair to say that approximately half of this chapter was written by Lars-Göran Johansson. The idea to write the article by evaluating string theory in terms of a number of different views on science and scientific methodology was originally my idea. The idea to address the difference between external and internal perspectives in the article was suggested by Lars-Göran. Besides this it is difficult to assess with hindsight, exactly who contributed with what in more detail, since we rewrote the article many times and had many discussions.
Chapter 5.
This chapter contains background material on different views on scientific theories in the philosophical discussion. The purpose of the chapter is to make the following chapters more readable, for people not very familiar with this topic. After presenting this background material, I also mention some of my own, more general views, concerning how we best can understand what scientific theories tell us about the world. This will then be further elaborated in following chapters.

Chapter 6.
This chapter is based on the article, (Matsubara forthcoming), which has been accepted for publication in, *Synthese*. The title of the article is, “Realism, underdetermination and string theory dualities”. This chapter focuses primarily on string theory. Some comments, on the relevance of the results to other approaches to quantum gravity, are however given.

Chapter 7.
In this chapter, the nature of spacetime and the claims, from string theory, that there exists extra dimensions, are discussed. An argument is developed to the effect that we cannot take the claims from string theory at face value. This chapter is based on a manuscript, which was written in collaboration with Lars-Göran Johansson. I have added a few reflections on the fundamental nature of spacetime, based on what is suggested by various approaches to quantum gravity.

Chapter 8.
This chapter is based on a manuscript I have been working on called, “The many worlds of physics and philosophy”. Various discussions in physics and philosophy, which all have that in common that they consider the existence of worlds or universes other than our own, are investigated. The main focus is on the discussion concerning the so called, “Landscape”, in string theory. This is then compared with other ideas such as, the many worlds interpretation of quantum mechanics and the discussion on possible worlds in philosophy. When possible worlds in philosophy are discussed I concentrate on the view of David Lewis; he thought of all possible worlds as equally real and existing in the same way as our own. A discussion on, how and when theoretical multiverse scenarios can be empirically testable, is also given.

\[4\] (Johansson & Matsubara submitted). The most important part of the argument, is based on considerations of the dualities in string theory, as described in the previous chapter. My contribution, to the manuscript, is primarily this argument based on the dualities. To avoid repetition, the background material about dualities in the original manuscript, is not included in this chapter.

\[5\] (Matsubara work in progress, a).
Chapter 9.
In this chapter I address the issue of thought experiments in quantum gravity. I investigate their potential use for the development of a theory of quantum gravity. My conclusions, concerning the use of thought experiments as an epistemic substitute for real experiments, are quite pessimistic. I do however see a use for thought experiments for research but in this context they can only be heuristic and it is difficult to see them as providing new reliable evidence.

Chapter 10.
This chapter contains a summary of the dissertation and some final comments and conclusions.

As has been described above, this text is to a large extent based on a number of articles and manuscripts, some of which I have written in collaboration with Lars-Göran Johansson. Of course the text, as it is presented in this dissertation, differs to some extent from the text in the articles and manuscripts. I have eliminated some parts so as to avoid unnecessary repetitions. I have also made quite a few other modifications to the text. I have for instance decided to change the pronoun "we" to "I", when it originally referred to me and Lars-Göran, in the texts we wrote in collaboration. I, after all, have to take full responsibility for the text as it now stands.

List of articles and manuscripts on which this text is based
2. Matsubara, K. (forthcoming). ‘Realism, underdetermination and string theory dualities’, *Synthese*. (The online version is accessible on PhilPapers.)

My interest, for the philosophical discussion on thought experiment, was initiated when I attended a course, given by Jim Brown, at the University of Toronto. This chapter is based on a manuscript, (Matsubara work in progress, b), which in turn is based on an essay I originally wrote for that course. In my manuscript, I have developed the argument made in the essay, and applied it more closely to questions concerning quantum gravity.
2. Physics and Philosophy

The attitude, among physicists, towards philosophy and its potential relevance to physics, of course differs from physicist to physicist. Some have a strong interest in philosophy and do find it relevant for their work. Despite this, it is not unfair to say that it is not uncommon among physicists to have a very negative attitude towards philosophy and philosophers. Is such an attitude warranted? Does philosophy have any relevance to physics?

In this chapter I will first focus on what a number of physicists have said about philosophy, exemplifying their attitude with a number of quotations. I have selected quotations expressing a negative attitude towards philosophy, without denying the existence of quite a few physicists having a more positive view on philosophy. After that I will give a defense for why and when philosophy is relevant for physics.

2.1 What physicists say about philosophy

A number of prominent physicists, have during the years, expressed negative attitudes towards the relevance of philosophy to physics. It is also not uncommon, among physicists, to utter negative remarks directed towards philosophers.

Especially well known is Richard Feynman’s rather negative attitude towards philosophers. A typical quotation is the following,

Philosophers say a great deal about what is absolutely necessary for science, and it is always, so far as one can see, rather naïve, probably wrong.

(Feynman 1963a, p. 2-7)

This quotation will be readdressed in Chapter 4. There it will be argued that physicists need to be quite careful when quoting Feynman as a critic of philosophy. For as can be seen there, even this critic of philosophy has said quite a few things regarding the very issues dealt with in philosophy of science.

Another example of a physicist, who has expressed negative attitudes toward philosophy, is Stephen Weinberg. In the book *Dreams of a final theory*, there is even a Chapter called “Against Philosophy”. However when reading the chapter, and the footnotes in it, it becomes apparent that it is rather specific views in philosophy that he is against. In a footnote he writes,

---

1See (Weinberg 1993, pp. 132-151)
Two philosopher friends have pointed out to me that this chapter’s title, ‘Against Philosophy,’ is an exaggeration, because I am not arguing against philosophy in general but only against the bad effects on science of philosophical doctrines like positivism and relativism...Anyway, I did not think that ‘Against Positivism and Relativism’ would be a very catchy title.

(Weinberg 1993, pp. 236-237)

Given this, I think it is fair to say that he does not really discard all of philosophy. Also, I would see the chapter itself, as arguing for a specific position in an ongoing debate in philosophy of science. To be against positivism and relativism is after all an attitude Weinberg shares with quite a few philosophers of science.

A quite recent example, of physicists criticizing philosophy, is found in the following quotation,

How can we understand the world in which we find ourselves? How does the universe behave? What is the nature of reality? Where did all this come from? Did the universe need a creator?...Traditionally these are questions for philosophy, but philosophy is dead. Philosophy has not kept up with modern developments in science, particularly physics. Scientists have become the bearers of the torch of discovery in our quest for knowledge.

(Hawking & Mlodinow 2010, p. 5)

However, while declaring that philosophy is dead, Hawking and collaborator did not hesitate to comment on the debate about realism concerning scientific theories. They introduce the term, “model-dependent realism”, writing that,

Model-dependent realism short-circuits all this argument and discussion between the realist and anti-realist school of thought. According to model-dependent realism, it is pointless to ask whether a model is real, only whether it agrees with observation.

(Hawking & Mlodinow 2010, pp. 45-46)

This sounds very similar to classical anti-realist instrumentalism; why is it thought of as a species of realism? The position is hardly original.² Here a famous scientist (together with not so famous co-writer) makes general claims criticizing philosophy as a whole, while still, quite sloppily, engage in that very discipline. When doing philosophy of physics it is important to keep up with progress in physics, but physicists talking about philosophy also need to keep up with the discussion in philosophy. I fall for the temptation to paraphrase Feynman and state that,

Physicists say a great deal about philosophy, and it is always, so far as one can see, rather naive, probably wrong.

²The debate about realism in science will be addressed in chapters 5 and 6. It is also relevant for the discussion in chapter 7.
This should not be taken too seriously though. Physicists have said ill informed, naïve and stupid things about philosophy just as philosophers have said ill informed, naïve and stupid things about physics. But it is equally true that not everything said by physicists about philosophy or philosophers about physics can be disregarded immediately.

One thing to keep in mind is that it is difficult to criticize philosophy, as a whole, for the following reason. Either you just disregard it without giving any real arguments; while this might be accepted by the ones already harboring a negative attitude towards philosophy, it will be less than persuasive for others. The other option is to give arguments, criticizing particular views in philosophy. However when doing that it is quite unavoidable to start doing philosophy and formulate your own philosophical view. In this way philosophy is unavoidable as long as there exists controversial issues to be discussed. Philosophy is not endorsing a specific view; talking about the view of the philosophers, in singular, is patently absurd.

Often I find that a physicist, who is dissatisfied with the views of a specific philosopher, could easily find an ally in another philosopher. Often one that has expressed critique of the first philosopher much better than what the physicist himself seems to be capable of.

My own view, on science and philosophy, is that there are no sharp boundaries between them. I also do not see the philosopher as an external judge of the practices of scientists, as might have been the view of some philosophers of science. I see it rather that the philosopher and the scientist can be engaged in a discussion concerning such topics as scientific methodology and how we are to best interpret and understand our scientific theories.

Regarding quantum gravity, theoretical physicists have for a long time been stuck in their attempts to connect their theories to observational reality. I find that this is not the best time for physicists to be arrogant. Rather it is more important than ever to consider philosophical and foundational questions concerning physics and its underlying assumptions.

2.2 Why philosophy is relevant for physics

Why then is philosophy relevant for physics? While I think that philosophy is relevant for physics, I do not want to exaggerate. Most of the time physics goes on quite well without the need of input from philosophy. However as stated above I see no sharp boundaries between physics and philosophy.

Much standard work in physics can be done without much philosophical thought. It is primarily when radically new theories are to be formulated or when physicists encounter controversial issues, where it is not easy to interpret or understand the theories, that philosophy is needed.

Regarding all approaches to quantum gravity, there have been a lot of criticism from various quarters. It is important that there are many critics among
physicists, it is not only philosophers who criticize the research done in these areas. When quantum gravity researchers need to defend their own views, for instance when critics among other physicists attack researchers in quantum gravity, they will themselves start to put forth arguments in foundational and philosophical issues. When doing that, I advise all of them to study what philosophers have said before. While some things might be particular to the research under discussion, there are other general considerations that philosophers might have investigated quite thoroughly. Questions for instance concerning the concepts used in theories or questions regarding how and to what extent we should trust what a theory tells us about reality.

On the other hand, I also find it important for the philosopher to keep up with modern science. I am quite critical towards much philosophy done at present, which in my view is too ill informed by modern science. With respect to these philosophers, the critique that they have not kept up with science is justified. However many philosophers of science are nowadays quite knowledgeable about the science they discuss.

I think that it is at the time, when new revolutionary theories are to be developed, that philosophical questions are most important. This can be compared with what Kuhn says about revolutionary periods in science, he argues that it is during these periods where philosophical questions become more important. There are quite a few things I disagree with Kuhn on but this is not one of them.

It is well known that in the early days of quantum mechanics the leading scientists, such as: Einstein, Bohr, Schrödinger and Heisenberg, did engage quite a lot in philosophical discussions. Also, it is well known, that they disagreed on many philosophical issues. This does however not mean that this discussion was useless or irrelevant to the development of physics. On the contrary it was a central part of the very development of quantum mechanics.

I think it is equally important that a new discussion on philosophical and foundational issues will be made concerning quantum gravity. And I find it that in that discussion it is important to have both philosophically informed physicists and philosophers with knowledge of physics, as participants.

---

3I am quite sympathetic towards much, if not all, of the criticism expressed in (Ladyman & Ross 2007). The authors of this book criticizes what they call “neo-scholastic metaphysics”. By this they mean; scientifically ill informed metaphysics produced by analytic philosophers. The philosophers engaged in what is judged to be neo-scholastic metaphysics, often have a very science friendly attitude. Unfortunately their lack of knowledge, about actual contemporary science, makes their cleverly thought out systems look quite implausible, from the point of a more scientifically informed reader.

4See (Kuhn 1970)

5For a few books, in which the various philosophical views of these physicists are discussed, see (Fine 1996), (Murdoch 1989), (Moore 1989) and (Heisenberg 1958).
3. Approaches to quantum gravity

In this chapter some background on the motivation behind the attempts to formulate a theory of quantum gravity will first be presented. This is followed by a short non-technical presentation of the two main approaches to quantum gravity. String theory will be given a slightly more detailed presentation than loop quantum gravity. Finally, I will briefly mention a number of approaches that are pursued by much fewer researchers, compared to the two main approaches. In that presentation I will also mention earlier research by pioneers in the field of quantum gravity, thus giving some hints on the historical development of the field.

3.1 Why do we need a theory of quantum gravity?

As has been mentioned above, it is not any empirical findings that motivate the search for a new theory. Instead there are two, very successful, theories of modern physics. On the one hand we have Einstein’s general theory of relativity on the other we have the standard model of particle physics which is formulated in terms of the language of quantum field theory.

While general relativity is quite different from Newtonian physics and introduces strange concepts, such as a curved spacetime to explain gravity, it is still a classical theory in the sense that it is not quantized. On the other hand quantum field theories are obviously quantum theories. The changes, to our view of the world, that was introduced by quantum physics were far greater and more radical compared to the changes from Newtonian physics that was brought about by the theories of relativity. And it is still the case that there are many issues that are not understood regarding quantum theory yet. There are a number of different interpretations and a lot of philosophical questions still worth to be discussed.¹

No discrepancies between these theories and the empirical results have been detected. They seem to work fine in the regimes where their predictions can be tested. So what is the problem?

¹This does not mean that there are no philosophical questions motivated by general relativity, see for instance (Malament 2007) and references therein. The interpretative issues regarding quantum mechanics are the most common topic of discussion in philosophy of physics; the literature on this topic is immense. Two overviews on this topic are (Dickson 2007) and (Wallace 2008). Much of the philosophical questions regarding quantum physics still deal with non-relativistic quantum mechanics, however recently more and more philosophical work has been done specifically on quantum field theory, see for instance (Brown 1988), (Teller 1995), (Auyang 1995), and (Healey 2007).
Unfortunately, the quantum field theory of the standard model cannot be combined with the general theory of relativity into a coherent theory. They are just too different. This wouldn’t cause much trouble for the physicist, in terms of predictions, if either quantum effects or gravity could be ignored in every possible situation. Often this is in fact possible, but not always; there are situations, for instance the conditions that obtain at the very early universe or in black holes, where both quantum effects and general relativity must be taken into account at the same time, and in such situations our ability to make predictions fail; we have no recipe for simultaneously applying the principles of quantum theory and those of relativity. Moreover, physicists seem to think that it is not satisfying with two so dissimilar theories about the physical world; there should be only one theory that accounts for everything.

In (Curiel 2001) it is argued that there really is not much of a conflict since we are not yet confronted with empirical results contradicting any theory. His view of theory formulation seems to be that theory in this case should follow empirical results. It would of course be beneficial to have some empirical data to start from, unfortunately that is not the case. Curiel argues that in previous situations when radically new theories was developed, such as the case with the development of quantum mechanics, the development was motivated by new empirical findings. In this case we do not have such a situation, instead it is primarily theoretical considerations that motivate physicists. This means that there are much fewer clues to go on and according to Curiel this also suggests that one ought to be quite pessimistic regarding our prospects of developing a successful theory of quantum gravity.

When developing a theory of quantum gravity it is quite probable that new ideas and concepts are needed. Without the aid of new experimental findings it is very difficult to find out what modifications to the old theories that should be made. Thus when trying to develop a quantum theory for gravity, based only on theoretical considerations, one must look at the theories that are already part of established science for clues. Of course the new theory must pass the empirical tests that the old theories have passed. The old theories ought ideally to be derivable as usable approximations of the new theory. The basic principles found in the established theories also provide clues. However some of these principles might not hold in the regimes where the new theory should be applied, they can not be seen as sacrosanct in all situations. Most approaches are quite conservative with respect to the basic principles. This might actually be a good strategy. In (Rovelli 2004, p. 6), it is pointed out that taking apparently conflicting theories seriously, can lead to major advances. One of the examples that Rovelli mentions is Einstein’s development of special relativity. It can be seen as taking Galilean relativity and Maxwell’s equations seriously and refuse to abandon both thus forcing him to instead radically modify our view on space and time. Different approaches to quantum gravity find dif-
ferent principles more or less important to hold on to when developing their respective theories.\(^2\)

### 3.2 String theory

String theory has developed tremendously since its inception around 1970.\(^3\) The basic idea is to explain the multitude of different particles as being different vibration states in quantized strings. It started off as a theory about the strong interaction. Originally a mathematical formula for certain scattering amplitudes satisfying a number of requirements, which were suggested by experiments and theoretical considerations, was found by Veneziano (1968), but there was no interpretation of this formula. The formula was later, by a number of researchers, derived using the assumption that they were dealing with quantized extended strings instead of point particles.\(^4\) The original version of string theory only contained bosons and needed 26 dimensions to be consistent.\(^5\) Since the real world contains fermions the theory needed to be modified. In this context the idea of supersymmetry was developed and this meant that the theory could deal with both bosons and fermions. It was then shown that a supersymmetric string theory require 10 dimensions in order to be consistent.\(^6\) This is an improvement from the requirement of 26 dimensions; still, it is more than the four dimensions in the observable world and hence an explanation for why the six additional ones are not observed was needed. It was imagined that the extra dimensions were very small and curled up or ‘compactified’.\(^7\) In later chapters of this dissertation, claims about extra dimensions will be scrutinized. It will be shown that the interpretation of these claims are not that straightforward.

\(^2\)For more on the general questions on why we need to find a quantum theory of gravity and the different approaches to it, see (Callender & Huggett 2001), (Rovelli 2007), (Rickles, et al. 2006), (Rickles 2008b), (Oriti 2009) and (Kiefer 2012).

\(^3\)Here I give a very brief description of this development. References to some important articles are also given. For the purpose of this dissertation it is not necessary to do justice to the finer points of the theory or its historical developments in any detail; only a short overview of string theory’s development, an overview that can function as background for the rest of the discussion, is required. To find more detailed descriptions and further references I refer to the vast literature on the subject. Standard textbooks are (Green, et al. 1987), (Polchinski 1998), (Zwiebach 2004) and (Becker, et al. 2007). Suitable accounts for the layperson written by professional researchers in string theory are (Greene 1999),(Greene 2004) and (Susskind 2005). The books by the critics (Smolin 2006b) and (Woit 2006) also contain easily accessible descriptions of string theory and its history. For a discussion on string theory written by a historian of science see (Galison 2007).


\(^5\)(Lovelace 1971).

\(^6\)(Neveu & Schwarz 1971) and (Ramond 1971).

\(^7\)There are other ways of explaining why we do not observe the extra dimension using so called D-branes. This is however of no importance for the arguments presented in this text.
In the meantime progress was made in conventional quantum field theory to
deal with the problems concerning the strong interaction; hence most of those
active in string theory abandoned the field and went to quantum field theory
instead. This eventually resulted in the formulation of the standard model.

String theory contains a spin two state that could be interpreted as a gravi-
ton, the quantum of gravity. This suggested that one could interpret string
theory differently, not as a theory for the strong interaction but as an all em-
bracing theory that could unify all fundamental forces and give us a working
theory of quantum gravity.\(^8\)

However, as long as progress was made towards the formulation of the stan-
dard model the interest in string theory was low and only few researchers ac-
tually pursued the idea. People also doubted that string theory could provide
a quantum theory of gravity free of the technical problems that had haunted
earlier proposals. But string theorists were able to give convincing arguments
for it being indeed possible to formulate a theory of quantum gravity, and after
this string theory became increasingly popular.\(^9\)

Soon it was found that five different supersymmetric versions of string the-
ory could be formulated, viz., type I, type IIA, type IIB, heterotic SO(2) and
heterotic \(E_8 \times E_8\). Today the generally accepted view, among string theorists,
is that these five versions of string theory are connected by dualities to each
other and to 11-D supergravity.\(^{10}\) They are all thought to be limiting cases of
a still more fundamental theory, called ‘M-theory’, which is still very poorly
understood.\(^{11}\)

It was also realized that string theory was not only about 1-dimensional
strings. Other many-dimensional objects called branes (a two-dimensional
brane is a membrane, thus the generic name) were also accepted to play an
integral part of the theory.\(^{12}\)

Strong arguments were then given that string theory allowed an immense
number of sufficiently stable vacua.\(^{13}\) This conclusion meant that the hope
to derive a unique low energy limit of string theory that reproduced the ob-
servable phenomena of our world was abandoned by many string theorists.
Susskind responded by introducing the term ‘Landscape’ to describe the set of
string theory solutions and was one of the early proponents of the use of an-
thropic arguments in the context of string theory.\(^{14}\) This issue will be further
discussed in chapter 8.

From this very brief overview it appears that string theory addresses the
following three main problems of theoretical physics.

---

\(^8\)(Scherk & Schwarz 1974) and (Yoneya 1973).
\(^9\)(Green & Schwarz 1984).
\(^{10}\)These dualities play an important part in the arguments put forth in chapters 6 and 7.
\(^{11}\)(Hull & Townsend 1995), (Townsend 1995) and (Witten 1995).
\(^{12}\)(Polchinski 1995).
\(^{14}\)(Susskind 2007) This was originally posted on arXiv already in 2003.
1. As has been mentioned above the four fundamental forces in nature are today understood in terms of two different types of theories, they are both very successful when applicable. The electromagnetic, weak and strong forces of nature are described by quantum field theory in terms of the standard model of particle physics, whereas the gravitational force is described in terms of Einstein’s general theory of relativity.\textsuperscript{15} As has been described the standard model and GTR does not work together. In string theory the goal is to find a complete unification of all the forces in one theory. This is a more ambitious goal than in other approaches to quantum gravity which are satisfied with providing us with just a quantum version of gravity. This problem, of finding a unified theory for all the four (known) forces, is the main motivation for string theory.

2. Another thing that concerns physicists is the large number of different particles and constants of nature. Their vast number make physicists wonder whether they really are fundamental and to look for a theory that explains them all as manifestations of something more fundamental. The string assumption provides just this.

3. A third problem is that the values of numerous parameters in the standard model are not derived theoretically from first principles. Instead their values are empirically determined without any fundamental explanation for why they have the values they have. For some this is not satisfying; they want an explanation of why these constants have precisely the values they have. Early on there was much hope that string theory could explain the values we observe. Since, nowadays, the standard view is that there are so many different solutions of string theory; the hope to derive unique values of the observed parameters has diminished considerably. On the other hand, there is in a sense, an explanation for how values of the parameters are decided by the theory.

For quite some time string theorists hoped that string theory would provide answers to these questions, but in recent years they have become more pessimistic, in particular about the possibility to derive unique values of the constants of nature. As stated above, it seems that there are many different models of string theory and that our universe is a realization of one of them. What this means for the question of the scientific status of string theory and its empirical testability is something that will be addressed later.

\textsuperscript{15}While the standard model and the general theory of relativity have been empirically very successful, the situation is not completely without problems. Straightforward estimates to calculate the cosmological constant based on quantum field theory give results that are way too high, by many orders of magnitude. This is a very interesting problem but will not be discussed further in this article. The interested reader can read more about the problem of the cosmological constant and different suggested solutions in for instance (Rugh & Zinkernagel 2002).
3.3 Loop quantum gravity

Loop quantum gravity (LQG) is the main contender to string theory, for solving the problem of quantum gravity.\(^\text{16}\) Still, it is a research programme, with much fewer active researchers than string theory.

While string theory, and its researchers, are more closely connected to the research tradition from particle physics, which developed the standard model, loop quantum gravity is more closely associated with the tradition that worked with GTR. Advocates of LQG often complain that string theory does not respect a fundamental, lesson learned from GTR, namely that a theory of spacetime ought to be background independent. In string theory what is done is to artificially treat part of spacetime as a fixed background and then do perturbative calculations around that background.

This is a central controversy between advocates of string theory and loop quantum gravity. On the one hand string theorists accept the critique but claim that the background dependence is only due to the present status of string theory. They believe that there exists a background independent formulation of string theory. Maybe a background independent formulation of string theory can be found when the mysterious M-theory is better understood. Also it is sometimes said that since string theory can be defined on different backgrounds it is already background independent. This is however not background independent in the same way as it is understood by defenders of loop quantum gravity. The issue of background independence is complicated and will be readdressed in chapter 7.

So then, what is the picture presented by loop quantum gravity? GTR can be formulated in a number of different ways, all equivalent at the classical level. As a starting point for LQG a Hamiltonian or canonical formulation of GTR is used.\(^\text{17}\) When a canonical formulation of GTR is used 4-dimensional spacetime is, in a sense, divided up again into 3-dimensional space and time. It is said that spacetime is foliated in terms of 3-dimensional spaces. This can be done in many ways, and at the classical level the end result will be the same. The reason for choosing a canonical formulation of a theory is that there is a quite straightforward and standard way to quantize such theories.\(^\text{18}\)

To straightforwardly quantize the most obvious canonical formulation of GTR does, for quite technical reasons, not work. Progress, that led to what be-

\(^{16}\)Standard textbooks on loop quantum gravity, containing references to original research, are (Rovelli 2004), (Thiemann 2007) and (Gambini & Pullin 2011). A popular science book describing LQG and other approaches to quantum gravity is, (Smolin 2000).

\(^{17}\)In physics ‘canonical’ is just a technical term used for theories described in the Hamiltonian fashion.

\(^{18}\)The technique I am talking about is the standard one, where the basic idea is to replace classical Poisson brackets with commutators of operators that act on a Hilbert space. This is either known by the reader, in which case further description is not needed, or not known by the reader in which case it can not be explained here.
came LQG, was only made possible after Ashtekar found a clever new choice of variables to describe GTR.  

Rovelli and Smolin used these new coordinates and found that the relevant quantum states could be described in terms of loops was related to loop states. It is unfortunately difficult to explain, in a non-technical fashion, what this really means. However the end result of this can be explained in a reasonably intuitive and pictorial way.

The idea is that the loops give us a description of space. It is not that the loop are embedded in space they are supposed to represent space itself. When loops might intersect there are nodes, these represent basic units of space. Space is hence discretized, two nodes connected with a link represent two units of space next to each other. The area of a surface is determined by the intersections with the loops. In this way a graph can be made with certain quantum numbers attached to it. The numbers determine areas of surfaces and volumes of space, this graph is a so-called spin-network. There are some conceptual issues with how to put the time back into the picture, this is the so called problem of time.  

A serious problem for LQG is that it has not been shown that it can reproduce GTR as a low energy limits. Scattering amplitudes have also not been calculated.

An approach closely related to, or even nowadays seen as part of, LQG is the use of so-called Spin-Foam Models. Spin-Foam Models, uses a path-integral approach to generate a spacetime. Roughly the development, in time, of spin-networks is supposed to represent spacetime, in terms of Spin-Foams.

3.4 Other approaches to quantum gravity

A large number of other approaches to quantum gravity has been put forth by various researchers. Here a few will be mentioned, without claiming that the list is exhaustive.

Early on, it was attempted to quantize the gravitational field in a way similar to what was done when quantizing other fields. However, when this is done, the metric of spacetime is artificially divided into a background, which

---

19 See (Ashtekar 1986) and (Ashtekar 1987).
20 See (Rovelli & Smolin 1988) and (Rovelli & Smolin 1990)
21 I will not here address the problem of time. The problem of time is discussed in for instance, (Rovelli 2007), (Rovelli 2004) and (Thiemann 2007).
22 Original articles are, (Barett & Crane 1998) and (Baez 1998). Spin-Foam Models are also discussed in (Rovelli 2004) and (Thiemann 2007).
23 Despite its close connection to LQG, which is based on canonical methods. Spin-Foam Models are covariant, just like string theory. In a way, Spin-Foam Models synthesizes different ingredients, from different approaches to quantum gravity, in an interesting way.
24 For more details see (Callender & Huggett 2001), (Rovelli 2007), (Rickles et al. 2006), (Rickles 2008b), (Oriti 2009) and (Kiefer 2012).
is treated classically. Only the deviations from this background are quantized. This violates an important principle of general relativity namely background independence; this is not very attractive. As has been mentioned above, this problem is also present in string theory. Even more problematic was that it turned out to be impossible to make such an approach workable, there were issues concerning non-renormalizable infinities etc.\textsuperscript{25}

There were also quite a few forerunners to LQG, in that they all used a canonical approach. As mentioned when LQG was discussed above it was quite difficult to make progress, using a canonical formulation, before the right choice of variables were made.\textsuperscript{26}

Path-integral methods, where sums were made over different geometries, using Feynman’s approach to quantum physics was worked originally worked on by Misner and developed by, among others, Hawking.\textsuperscript{27}

Another approach is connected with Penrose’s twistors.\textsuperscript{28} This is a reformulation of geometry, which can be done at the classical level. The hope was that the reformulated version would be more suitable for quantization. While a number of interesting mathematical results have come from this, no real success has been made with the project of formulating a theory of quantum gravity.

In ordinary quantum mechanics, the phase space, is noncommutative. This can be described in terms of a so called noncommutative geometry. When this is used, as part of an approach to quantum gravity, it is assumed that the same would be the case for spacetime itself.\textsuperscript{29}

Interestingly, mathematical concepts such as twistors and noncommutative, geometry have appeared in the context of string theory.\textsuperscript{30} It might be that twistors or noncommutative geometry are not by themselves sufficient, to build a whole approach to quantum gravity, while still being part of a bigger picture.

Another approach is causal set theory. Here a discrete model of spacetime is built by using the mathematical theory of causal sets. The points are partially ordered, in a way representing their causal relations.\textsuperscript{31} According to (Rovelli 2007, p. 1296), this approach actually predicted a small but non-vanishing cosmological constant. The value was also of the correct order of magnitude compared to the recently observed experimental data. Hence, if this is true,

\textsuperscript{25}A few important original research papers, in this approach, are in chronological order (Rosenfeld 1930), (Fierz & Pauli 1939), (Gupta 1952), (Feynman 1963b) and (Deser & van Nieuwenhuizen 1974).
\textsuperscript{26}A few early papers are (Bergmann 1949), (Dirac 1950), (Peres 1962), (Arnowitt, et al. 1962), and (De Witt 1967).
\textsuperscript{27}(Misner 1957), (Hawking 1978).
\textsuperscript{28}(Penrose 1967).
\textsuperscript{29}For more on this see (Connes 1990) and (Madore 1999).
\textsuperscript{30}See for instance (Seiberg & Witten 1999) and (Witten 2004).
\textsuperscript{31}(Sorkin 1983).
this would be the only real confirmed prediction from an approach to quantum
gravity.\textsuperscript{32} I guess the lack of real enthusiasm, over this result, is due to the
fact that the approach to quantum gravity based on causal sets is not that well
developed and that there are quite a few problems with the approach.

As can be seen above, there are quite a few different approaches to quan-
tum gravity. It might be the case that, in some clever way, combine insights
from various approaches is the right way to formulate a successful theory of
quantum gravity.

\textsuperscript{32}I must admit that I have not studied the original text where this prediction is made, I have only
seen the comment by Rovelli.
4. Methodology and research in quantum gravity

This chapter is closely based on (Johansson & Matsubara 2011), in which the focus was on string theory. To broaden the scope, in 4.4, I have added a few comments regarding loop quantum gravity and other approaches to quantum gravity.

4.1 String theory

String theory has evolved into a dominating field of research, perhaps the dominating one, in fundamental theoretical physics during the last 30 years or so. For a long time the hope among those in the field was that it should be The Theory of Everything, that it should give us the answers to all the remaining problems in fundamental physics.\(^1\) But the hope has so far not been fulfilled and signs of increasing uneasiness in the theoretical physics community can be seen. There has always been some opposition against the theory and in recent years it has gained increased strength. The critics that probably have received most attention are Smolin and Woit. In The Trouble with Physics, (Smolin 2006b), and Not Even Wrong, (Woit 2006), string theory’s dominating place in theoretical physics is questioned. Their main argument is string theory’s lack of new testable predictions despite heavy efforts by a huge number of devoted physicists. But the majority of string theorists do not seem deeply concerned; most still seem to be in good mood, although the attack by Woit and Smolin has not passed unnoticed.\(^2\)

A first rough assessment of the situation is this: the majority of string theorists are not convinced by the criticism, but they feel a need to come up with a defense. This is profitable for the philosopher, because such a debate will stimulate people to formulate hitherto more or less tacit assumptions and norms. Obviously, those involved do not think the present lack of testability is reason enough to give up. String theorists have not abandoned the quest for testable results, as judged from what they say and write; on the contrary, it is still eagerly wanted. But the situation is problematic and warrants further discussion.

\(^1\)The world ‘everything’ should not be taken literally; everything physical is the intended scope!

\(^2\)Cf. for example (Chalmers 2007).
In this chapter it will be discussed what philosophy of science can contribute to the debate concerning the scientific status of string theory. Earlier discussions on string theory have often become too polarized and often too simplified. Here an attempt is made to give a more nuanced discussion and take a look at string theory from a number of different perspectives on science. What conclusions could be drawn from the perspectives of, respectively, logical positivism, Popper, Kuhn and Lakatos? For a person with a strong allegiance to a specific position in philosophy of science, this might facilitate a decision on how to evaluate contemporary research in quantum gravity.

String theorists hope that their efforts will result in an all embracing theory, a theory that explains everything. Even a superficial reading indicates that the explanation wanted is connected to realistic attitudes, whereas some critics have less metaphysical convictions. Hence the realism/anti-realism issue is connected to questions about form of explanation and explanatory value.

An important point is that one must distinguish between the mathematical framework of string theory and its physical applications. In this text it is assumed that string theory is a theory attempting to formulate a quantum theory of gravity and unify the fundamental forces, the single exception being when it is described how string theory was originally thought of as describing strong interactions. Applications of the mathematics of string theory to other areas is not considered to be what I mean by “string theory” in this chapter.

Another topic that is relevant is whether to adopt an externalist or an internalist perspective on scientific change. The author’s own views on these questions will be revealed later in this chapter, but the main purpose here is to illustrate how different assumptions regarding science lead to different evaluations of the scientific status of string theory. Most of these conclusions apply to other approaches to quantum gravity as well.

Theoretical physics asks the most profound and general questions about physical nature such as: What are the ultimate constituents of the universe? Is there one unified theory of everything? How did the universe begin? Is there any explanation of Big Bang? String theory aims to answer some of these questions.

Such questions are intimately connected to philosophical issues. For example, what kind of explanation do physicists have in mind? Why is a unified theory of everything desirable? What norms are used when deciding which is the best alternative route to develop an area of research? What needs to be explained and what can be taken for granted, not requiring explanation?

In this chapter I will address some of these philosophical issues; ontological, epistemological and methodological assumptions made by the string theory community will be described and assessed. Criticism directed towards research in string theory will also be evaluated. As has already been stated, this will be done, not by just taking one view on science and scientific method for granted when considering the scientific status of string theory, but to discuss the questions from different perspectives.
4.2 String theory and philosophy of science

Four well-known theories regarding the development of science are: logical positivism, Popper’s falsificationism, Kuhn’s theory of scientific revolutions and Lakatos’ theory of scientific research programmes. I will here give very brief descriptions of these views and discuss what an adherent to the respective view would say about the present state of string theory.

4.2.1 String theory from a logical positivist’s perspective

The logical positivists view on science have been severely criticized and it is not a popular account of science today. Nevertheless the view has been very influential in past decades and some of the ideas supported by the logical positivists are still influential on how scientists view their theories. It should be noted that the logical positivists considerably developed their views during its heydays and there was no consensus on all issues, so the presentation here is indeed simplified. The main goal for logical positivists was to eliminate metaphysical discussions from science and secure a firm empirical basis for the sciences. To achieve this goal they proposed a criterion for empirical meaning of any statement, viz., that it should be given in terms of the methods for verifying it. If no such verification method could be given it had no meaning, i.e. no truth value. The main problem is to apply this criterion to theoretical statements. Their solution and final view was that laws, and theoretical statements in general, lack truth value; only observation statements have a definite truth value and can be verified. Laws was held to be mere instruments for predictions. This of course presupposes a very sharp distinction between observation statements and theoretical statements and this proved to be one of the profound and unsolved problems of logical positivism.

Logical positivism could be characterized as consisting of verificationism in semantics, inductivism in methodology and an instrumentalist view on theories.

Now, what would a logical positivist say about string theory? Would he repudiate it tout court as speculative metaphysics?

First, the instrumentalist aspect of logical positivism doesn’t accord well with string theorist’s own views. Since an instrumentalist holds that theoretical statements are tools for predictions, he has little reason to ask for a unification of the four interactions, which seems to be the ultimate goal for many string theorists. Since there are no observations disagreeing with our existing theories, asking for unification is most naturally interpreted as expressing a realist attitude towards theories, since the realist has the ambition to describe an independent reality behind the observations. In contrast an instrumentalist would be satisfied with any set of assumptions, postulates or hypotheses that

---

3 See (Ayer 1952), (Ayer 1959),(Carnap 1966), (Nagel 1961) and references therein. A detailed discussion on logical positivism can be found in (Suppe 1977).
make correct predictions of observable events. At most, a logical positivist could endorse the search for unification for pragmatic reasons, if that would make calculations and predictions easier.

Second, work in string theory is anything but an example of inductivist methodology, it is rather more like suggesting new explanatory hypotheses; a number of additional assumptions are introduced by which one is able to reproduce quantum gravity and embed quantum field theories of the types that appear in the standard model. But since no new testable predictions hitherto has been made, one cannot say the new hypotheses have inductive support. According to logical positivism, observations ideally comes first and theory construction is seen as a systematization of empirical data. Observation and theory construction is supposed to go hand in hand according to logical positivists, which is certainly not the case in string theory.

Third, it seems impossible to interpret string theory as encapsulating semantic verificationism. However, this last point is in our view of very little interest, since few nowadays accept the form of verificationism espoused by logical positivists.

Would a logical positivist dismiss string theory as metaphysical speculation? An adherent of the first version of logical positivism, which required complete reduction of theoretical statements to observational statements, would do so, since string theory fails the verification criterion. A defense claiming that string theory is verifiable in principle, using large enough accelerators, doesn’t sit well with positivism; it is only when we actually can put our theory in direct connection with empirical observations that the theory can be said to fulfill the verification criterion. But an adherent to the later version, where theoretical statements are viewed merely as a device for calculations, could be less hostile. If string theory would succeed in reproducing all “low energy data” hitherto accounted for by GTR and the standard model, it would have the same status as these two theories, viz., tools for calculation. But so long as no new testable predictions has been made, this tool is no better than the earlier ones.

Summarizing, from a logical positivist’s point of view, the basic assumptions in string theory are neither true nor false, they are just tools for making predictions about observable phenomena, and since no improvements in this respect have been made, string theory is unsuccessful.

4.2.2 String theory from a Popperian perspective

To begin, an important aspect of Popper’s views on science was to sharply distinguish between context of discovery and context of justification. The importance of this distinction is that within the context of discovery there are

---

4The standard text is (Popper 1959).
5This was however a view that was also endorsed by the logical positivists.
no methodological rules; the scientist can use whatever means he finds suitable, including metaphysical speculations, when trying to formulate interesting hypotheses. In contrast, Popper had very strict norms for the testing of hypotheses. He endorsed the hypothetico-deductive method, albeit with some idiosyncrasies of his own.

Popper’s basic norm is that no theory that cannot be falsified may be called scientific. Second, he holds that the scientist should reject a falsified hypothesis and invent a new one. His third norm is that ad hoc assumptions should be avoided. If a new assumption is introduced, to save a theory from falsification, it must be independently testable according to Popper’s view on scientific method.

Now let’s move to string theory; how would a Popperian evaluate its present state? Has string theory been put to a test yet and has it entered a phase of ad hoc excuses, in either the stronger or the weaker sense, that eliminate the testability of the theory?

From a Popperian perspective it may seem ad hoc to say that there are a number of extra dimensions that are compactified. Defenders might say that extra dimensions are not introduced as ad hoc assumptions, they are a core idea in string theory. But, a Popperian may ask “why compactify precisely six dimensions?” This idea is obviously geared so as to fit our observations of a four-dimensional spacetime.\(^6\) The crucial thing is whether compactification of precisely six dimensions is in principle independently testable, which is the reasonable demarcation for ad hocness. This question is, so far as one can see, presently undecidable and therefore one could not rule out that this assumption could be tested in the future. On the other hand, there is a reasonable suspicion that the assumption of compactification of exactly 6 (or 7, if we start from an 11D M-theory) dimensions could be seen as ad hoc; why not all 10, or just one, as Feynman asked in an interview (Davies & Brown 1988, p. 194).

Presently, string theory is not testable in the sense that no new testable consequences, not already implied by the standard model or GTR has been derived; this is the basic criticism against it and it is pretty clear that the critics are inspired by Popper’s views.

But people in the field, Susskind for example, are not impressed. Susskind calls the critics ‘Popperazzi’ and continued by quoting Feynman:

> Philosophers say a great deal about what is absolutely necessary for science, and it is always, so far as one can see, rather naïve, probably wrong.

(Susskind 2005, p. 192), quote taken from (Feynman 1963a, p. 2-7).

Susskind thus thinks that criticism inspired by Popper is not to be taken seriously and he believes to have an ally in Feynman on this point. But he has not,

---

\(^6\)In chapters 6 and 7, the claims from string theory, that there are extra dimensions of spacetime, will be critically examined. And it will be seen that it is not so straightforward to interpret the formalism of string theory.
for Feynman’s view in methodology is more or less text-book Popperianism, as can be seen from the following quotation:

In general we look for a new law by the following process. First we guess it. Then we compute the consequences of the guess to see what would be implied if this law that we guessed is right. Then we compare the results of the computation to nature, with experiment and experience, compare it directly to see if it works. If it disagrees with experiment it is wrong. In that simple statement is the key to science. It does not make any difference how beautiful your guess is. It does not make any difference how smart you are, who made the guess, or what his name is - if it disagrees with experiment it is wrong. That is all there is to it.

..... You can see, of course, that with this method we can attempt to disprove any definite theory. If we have a definite theory, a real guess, from which we can conveniently compute consequences, which can be compared with experiment, then in principle we can get rid of any theory. There is always the possibility of proving any definite theory wrong; but notice that we can never prove it right. Suppose that you invent a good guess, calculate the consequences and discover every time that the consequences you have calculated agree with experiment. The theory is then right? No, it is simply not proved wrong. In the future you could compute a wider range of consequences, there could be a wider range of experiments, and you might then discover that the thing is wrong.


It is also known that Feynman was quite skeptical about string theory, as revealed in an interview Mehra made with him (Mehra 1994, p. 507).

Leaving aside the views of Feynman, how did Susskind defend string theory against criticism from a Popperian perspective? His strategy is argument by historical analogy. He gives a number of examples from the history of science in which a hypothesis that once appeared impossible to test now is not only tested but also universally accepted. Susskind concludes:

Good scientific methodology is not an abstract set of rules dictated by philosophers. It is conditioned by, and determined by, the science itself and the scientists who create science.

(Susskind 2005, p. 192)

I agree with Susskind’s general remark here, that rules cannot be settled once and for all by armchair philosophy. But it is equally important to notice that Susskind does not question the requirement of testability, only the, in his opinion premature, conclusion that a particular hypothesis is not in principle testable. So in defending string theory against the critics Susskind makes an inductive move; we have seen many cases of theories that once appeared not

7The interview was conducted shortly before Feynman’s death in 1988. While a lot of new developments have occurred in string theory since then, I think there is little reason to think that Feynman would have changed his views on string theory, given his general attitude towards physics.
testable but later has been possible to test, hence we are well advised to wait and hope that also string theory one day will admit testing.

Is it, then, from a Popperian perspective, correct to say that string theory is not acceptable science? This is far from clear. Popper himself held that wildly speculative assumptions, even metaphysical ones, are admissible in science, if they help develop testable hypotheses. Furthermore, he thought that what is required is falsifiability in principle, not that it must be possible to falsify a hypothesis immediately after its formulation, or at any particular moment of time. So a Popperian could accept string theory as legitimate research while stressing that it has not (yet) produced any testable hypotheses. The crucial thing is, of course, whether it’s lacking testable implications really is a momentary or permanent feature. In contrast to the positivists, in principle arguments are more palatable for Popper, the reason being that he did not tie his views on scientific method to any semantic theory. So presently unfalsifiable claims are considered meaningful and this accords with the view among string theorists themselves. The most reasonable description of string theory from a Popperian perspective is that it is a viable project to work on, despite it being not yet testable. However it is likely that a Popperian would say that although it is a legitimate area of research it has not yet generated a truly scientific theory. A Popperian would also stress that it is important to be open with this and not claim success prematurely. It seems that Susskind could accept that.

The critics think, of course, that the time to give up is overdue, while the defenders say it is not. Popper’s methodology, interpreted as “in principle falsifiability” cannot decide between them.

4.2.3 String theory from a Kuhnian perspective

Kuhn was much more descriptive and less normative in his account of science than Popper and the positivists. Furthermore, he adopted a third-person, i.e., externalist perspective in philosophy of science, which entails that scientific changes not always could be explained using the model of rational deliberation as applied to individual or collective decisions in the scientific community. Sometimes it is not scientific reasons that ultimately explain theoretical change; instead social and other external causes are invoked. His main reason for this stance is that the conditions for rational theory choice is not always satisfied. Kuhn describes the development of science as alternations between periods of normal science and periods of revolutionary science.

Kuhn considers the periods of normal science to be of more importance than what would be suggested by a Popperian account, which focuses on

---

8(Kuhn 1970).

9While Kuhn is clearly more of an externalist than Popper and the logical positivists, it is interesting to note that he nevertheless saw himself primarily as an internalist, as can be seen in (Kuhn 2000, pp. 286–287).
falsification and the refutation of previously held theories. Kuhn described scientific work within a discipline during a normal period as governed by a paradigm consisting of a set of metaphysical assumptions, formalisms, norms and paradigmatic exemplars of successful solutions to scientific problems. Normally these beliefs and habits are hardly articulated. Failures are usually not seen as evidence against the paradigm; they are quarantined for the time being as anomalies. Scientists working within the paradigm assume that these anomalies can be dissolved by modifications of auxiliary assumptions while the basic convictions are left unchanged. According to Kuhn it is these periods of normal science that is the hallmark of science. It is when you take a number of fundamental assumptions for granted and work with ‘puzzles’, which you try to solve in light of these assumptions, that most scientific results are accumulated. If the scientists constantly were debating fundamental questions they would not make any progress. According to this view it is perfectly OK to use auxiliary assumptions to protect the basic assumptions of the theory. But it might happen that anomalies accrue, scientists begin to feel the pressure of many unsolved problems and they start reflecting upon their basic assumptions previously taken for granted. A crisis is on the way, and this may finally result in a replacement of the ruling paradigm for another one.

Such a change is labeled ‘paradigm shift’ or ‘scientific revolution’; it is a change from one period of normal science to another. Kuhn does not describe this change as the result of a rational choice made by the discipline, since the condition for rational choice, comparability with respect to goal achievement, or commensurability, is not fulfilled. In short, paradigms are, according to Kuhn incommensurable.

Many philosophers are highly critical of Kuhn for his incommensurability thesis and we join the critics when it is directed to some extreme interpretations of incommensurability, viz., those that entail complete relativism. But I need not take any stance on that matter at this point. Be as it may, let’s look upon string theory with Kuhnian glasses and ask ourselves: is the present status of string theory that of a normal science? Is there anything like a string paradigm? Yes, I think so; there are symbolic generalization, metaphysical assumptions values and exemplars, which upholds a puzzle-solving tradition.

---

10 As can be read from the many essays published in (Kuhn 2000), Kuhn himself did not regard his views as expressing relativism. He argued that science does progress, in an objective sense. However he did not believe that one was justified to claim that our theories converged towards better and better descriptions of an underlying reality. In this sense he was a scientific antirealist.

11 It is claimed in (Audretsch 1981) that Kuhn’s theory doesn’t fit theoretical physics, because there are two paradigms in it, a quantum field paradigm and a geometrical paradigm, the latter being the framework for GTR. Each paradigm completely dominates its domain of application. I don’t agree with Audretsch’ conclusion; for Kuhn, discipline borders are described in terms of paradigms, so if there are two non-competing paradigms, they define two sub-disciplines. But I agree with his point that quantum gravity may be seen as an effort to unify two paradigms, a situation which is not discussed by Kuhn.
In addition, while there are competitors, string theory has a very dominant status, which is a criterion for being normal science. The odd thing is that it acquired this status before it has been able to make any successful empirical predictions.

The next question is if one could use Kuhn’s theory to decide whether string theory is in crisis and ripe for dismissal?

My answer is a qualified no. Kuhn’s theory was never aimed at giving a methodology with normative force, not even a recipe for determining the status of current paradigms. It is best viewed as an account of history of science, made in retrospect and from an externalist perspective. Its aim is not to give any effective and usable criteria for when a paradigm will crumble, but rather to describe the historical process in retrospect.

Are there more and more anomalies? Is there a feeling of crisis among string theorists? The answer is presumably no. There have been profound unsolved difficulties all along, of which some are solved, some not, and their number has, as far as I can see, not obviously increased. That some critics have voiced their opinions the last years is a sign of increasing dissatisfaction, of course, but the majority doesn’t seem deeply disturbed. Thus, so far no revolution is in sight. But one should keep in mind that that could change rapidly and Kuhn’s theory doesn’t help us by providing any predictions. Kuhn’s theory is, to repeat, basically descriptive and historical.

4.2.4 String theory from a Lakatosian perspective

At first sight, one may say that Lakatos tried to steer a middle way between Kuhn’s non-rational and non-normative account of paradigm change and Popper’s strongly normative falsificationism. In his seminal paper, (Lakatos 1970), he introduced what he called the Methodology of Scientific Research Programmes, MSRP for short. It contains a methodological norm, or so it seems, while at the same time it takes into account the feature observed by Kuhn that scientists seldom give up a theory or hypothesis when confronting a few unfavorable test outcomes.

One important difference between Kuhn and Lakatos is that Lakatos held, whereas Kuhn denied, that research programmes can be rationally compared; they are either degenerative or progressive and the scientific community, if rational, chooses the progressive over the degenerative one.

Another important difference between Popper and Lakatos is that Lakatos accepted, while Popper denied, that there are crucial experiments that may confirm a theory. These may be conducted when we have a new research programme as an alternative to an old one; Lakatos maintained that in a crucial experiment, one hypothesis is rejected and the competing one is confirmed. He also differed from Popper in holding that a failed test is by itself not sufficient reason to give up; Lakatos stresses that we do not, and should not, reject a
research programme if there are no alternatives in sight; better to work on the problems in an unsuccessful programme than to simply give up.

Lakatos defined a research programme as consisting of a series of theories and characterized by four components: i) the hard core, ii) the protective belt, iii) the positive heuristic and iv) the negative heuristic. The hard core comprises those assumptions that are common to all theories in the series. That which changes from one theory to another belonging to the same programme is the protective belt. Positive and negative heuristic tells the researcher what to do, and what not to do, when pursuing the research programme.

Lakatos overarching methodological rule for change of research programmes was: replace a degenerative research programme by a progressive one! The distinction between degenerative and progressive research programmes is made in terms of what comes first, theoretical or empirical development. In a progressive programme, theoretical advancements suggest new experiments that then are planned and conducted. Then the empirical results come out most often as predicted. This is to be contrasted with a degenerative programme where empirical findings force the scientists to modify their theory into newer versions within the programme by replacing components in the protective belt. In a degenerative research programme the function of the protective belt is akin to that of ad hoc assumptions; the difference is that independent testability might be possible. So the methodological rule to give up a degenerative research programme has an effect analogous to Popper’s rule not to accept ad hoc assumptions.

But when has a research programme entered a phase of degeneration so it is time to give up? Even a very progressive programme could come into a degenerative phase for some time, and scientists were by Lakatos advised not to give up too hastily. It is often rational to wait for some time before giving up. In fact, Lakatos has not given us any effective methodological rule by which we could arbitrate between rational and irrational actions. Both to give up and to stick to an ongoing research programme could be defended as rational, given Lakatos’ prescriptions. Unfortunately, Lakatos died before he had answered this criticism, so it is not known what he thought about it. Obviously, Lakatos theory needs improvement.

Or does it? Perhaps what is needed is only a new perspective to evade the criticism? This was Hacking’s point of departure in his assessment of Lakatos’ theory in (Hacking 1983). Hacking’s defense was that Lakatos could be interpreted so as to not give methodological prescriptions for the individual scientist, but to give a method for rational reconstructions of periods of history of science. The development of a science during a period can be judged rational or not rational by the standards chosen, post hoc, by the philosopher or historian of science, independently of the beliefs of those working in the evaluated discipline.

Thus, the advice to give up a degenerative research programme in favour of a progressive one might not be an advice directed to scientists but an advice
directed to the historian aiming at rational reconstruction of science history. Every historian has to choose a perspective and some principles for selecting which facts to take into account and which to neglect. Thus Lakatos methodological rule is better stated as: reconstruct periods of history of science in such as way that scientists are depicted as normally giving up degenerative programmes in favour of progressive ones when the objective conditions are fulfilled, disregarding their beliefs.

It does not follow, from adopting Hacking’s interpretation of MSRP, that the concerned scientists, the agents being described, are not given any methodological rules. Knowing about MSRP in Hacking’s interpretation, accepting it and wanting to be rational, they should certainly follow the replacement rule if possible. But they have not enough information to definitely judge their alternatives, so the replacement rule is not an effective one for them. It is only effective for the historian.

According to a Lakatosian view one may also accept that many research programmes can exist in parallel and adherents to different programmes may very well agree that one programme is more progressive than the other at any particular point of time. But there is nothing irrational to pursue a less progressive programme, although it appears that the risk for failure is higher. One has no argument for saying that minimizing risks in choosing research programme is the most rational option.

If one research programme is particularly progressive it will presumably attract most researchers in the field and other programmes are simply abandoned. This is natural if it is believed that scientists share a minimum set of values, attitudes and vaguely defined norms.

After these interpretative considerations it is time to take a look at string theory as an instantiation of a Lakatosian research programme. Here is a selection as to what could be a plausible reconstruction of string theory’s content, or part of it, in terms of the four components of a Lakatosian Scientific Research Programme:

- **Hard core:**
  i) The fundamental objects are not point particles but extended objects, strings or branes.
  ii) Accept the basic assumptions of quantum mechanics as given.
  iii) Require supersymmetry of the theory.\(^{12}\)

- **Protective belt:**
  i) The different versions of string theory are merely different theory formulations, not different theories.
  ii) Compactified dimensions are too small to be observed with present day technology.

\(^{12}\)We assume that we talk about superstrings here. It should also be noted that supersymmetry should be broken at lower energies in order to give a correct description of our world.
iii) Explain the value of the constants of nature assuming a landscape of universes.

- Positive heuristic: Develop the theory so one can:
  i) explain the diversity of particles as mere manifestations of one fundamental kind of objects.
  ii) derive the constants of nature, iii) unite the standard model with gravitation.

- Negative heuristic: Don’t allow any modus tollens argument to be directed against the hard core.\(^{13}\)

Comparing string theory with its competitors, e.g. loop quantum gravity, it seems clear that the string programme has attracted most researchers, which have developed the theory in many respects, and so one could say it has been progressive in a more general sense.\(^{14}\) On the other hand, using Lakatos criteria for distinguishing between progressive/degenerative ones, no such divisions can be made, simply because empirical tests has been lacking.

However, there is a sense in which string theory confronts the observable world. Take for example constants of nature. We have known their values for some time. String theorists during one period hoped to be able to derive them from first principles. The derivation of the actual constants failed, one got a vast number of possible values and combinations. In response some came up with the idea of a multiverse, consisting of an enormous set of universes, each characterized by a unique combination of values of those constants. This move fits rather well with Lakatos’ description of a degenerative phase in which empirical findings drive the theoretical development, although in this case the empirical results were known in advance.

Is the difference between knowing the empirical results in advance or not relevant for the question whether the programme is degenerative? Of course it is.; in a truly progressive programme new empirical results are predicted. But conflicts with previous experiments could still be considered to have some relevance. The move of postulating a multitude of universes as a response to a failed prediction/retrodiction could plausibly be described as an instance of a degenerative phase of the research programme. For there is nothing in the hard core of the research programme indicating a multitude of universes, nor that that assumption is needed for protecting the hard core from being targeted by a modus tollens argument. This does of course not mean that the hypothesis that there are multiple universes necessarily is wrong, but one should at least be a bit worried. The existence of these multiple universes

---

\(^{13}\)This means that if observations do not agree with predictions we should try to achieve consistency between predictions and observations by modifying or adding assumptions in the protective belt. The assumptions in the hard core should never be blamed.

\(^{14}\)It should be noted that quantum loop gravity while being a contender when it comes to quantizing gravity do not attempt to unify all forces. Despite this less ambitious goal I think it is fair to describe quantum loop gravity as a competitor to string theory.
String theory is a degenerative programme, according to Lakatos’ criterion; the empirical facts against which string theory is tested was known in advance, no new testable empirical predictions have been made, and the mismatches that have been found have been a driving force in its development. Hence, if Lakatos is read as proposing a decision rule, the conditions for dismissing a research programme would be fulfilled, if there was a progressive rival programme. This is however not the case, so Lakatos’ criteria cannot really be applied, because the rival programmes are not more empirically successful than string theory.

4.3 Further remarks concerning string theory and scientific norms

There is no general agreement about which perspective to adopt in general methodology; hence, the evaluation of the scientific status of string theory must be conditional on which stance one takes. Nevertheless, some general conclusions are possible. The first is that, except for adherents to a strict verificationism typical of the early logical positivists, no matter which of the discussed positions you take, you cannot definitely and clearly reject string theory as unscientific. Second, again with the exception of the verificationist, no methodologist could definitely tell string theorists that it is time to give up and do something else. Perhaps somewhat surprisingly, not even a convinced Popperian could say that string theory should now be rejected as a falsified or unfalsifiable theory.

In my own view, the failure of now making a definite judgment about what is rational vis a vis string theory is not at all astonishing. For if a methodologist were able to do that, he would in this particular situation have solved the induction problem; he would be able to tell us, on the available evidence from history of science, what is the correct inference, i.e., he would be able to tell us what is the rational way to proceed. But I don’t think that is possible; as Quine once wrote ‘The Humean predicament is the human predicament’ (Quine 1969, p. 72).

I hold Lakatos theory, MSRP, to be the most reasonable analysis of scientific development; it fits quite a number of episodes from history of science and I think it strikes the right balance between a descriptive and a normative account of science. It is also, to some extent, useful for discussing string theory and its competitors, mainly loop quantum gravity. However one cannot really say that one programme is progressive and one degenerative, because the distinction and comparison is made in terms of theoretical and empirical development, and no empirical development has occurred. On the other hand, without using Lakatos criteria and instead merely relying on our somewhat
vague notion of development, one is tempted to say that string theory has been *theoretically progressive*, but not *empirically progressive*. One could say that adherents to string theory believe that theoretical progressiveness is sufficient for continuing work on the theory, whereas critics think it’s not.

In a short article, (Cartwright & Frigg 2007), conclusions somewhat similar to the ones given in this chapter are presented. The authors end up by evaluating string theory in terms of Lakatosian research programmes, although they have used a broader set of criteria for determining string theory’s degree of progressiveness than Lakatos did. They mention such virtues as i) a large range of empirical applications, ii) successful novel predictions, iii) spawning new technologies, iv) answering perplexing problems, v) consistency, vi) elegance, vii) explanatory power, viii) unifying power, and ix) truth. They conclude that string theory has been progressive along the dimensions of explanatory and unifying power, but that is not sufficient for saying that string theory is generally progressive. The authors however does not recommend dismissing string theory; referring to Lakatos, who is said to recommend us to “treat budding programmes leniently” they conclude that string theory deserves still being pursued. I agree that Lakatos MSRP does not give a definite recommendation in a case like the present one. I also agree that string theory deserves still being pursued, albeit I see no good reason why most efforts in this area should be put on improving string theory; it isn’t obviously superior to its competitors.

Their list of dimensions of progressiveness is hardly an interpretation of Lakatos’ notion of progress, it is their own version. Two of their criteria, elegance and truth, deserve a comment. As regards elegance, it could reasonably be held that elegance, however it is made precise, is not a property of the object but is in the eye of the beholder; it has to do with scientific taste and other non-epistemic aspects and I don’t think such criteria should be used. I also question their mention of truth as one dimension along which a research programme should be evaluated. It can be inferred that a certain theory plausibly is true, or approximately so, using all available evidence. Thus truth of a theory is not a criterion but the aim.

One important aspect, of Lakatos MSRP, is not stressed in (Cartwright & Frigg 2007), viz., that one should abandon a degenerative research programme in favour of a progressive one. In quantum gravity there are no truly empirically progressive competitors to string theory, hence this rule can not be applied. So what is the suitable strategy to use when no programme is progressive? Neither Cartwright and Frigg, nor Lakatos give a definite answer; some people may take the risk on working with alternative research programmes before they reach the progressive state, some others may not, and neither choice seems irrational. I return to this point at the end of this chapter.

I think that Kuhn has widely exaggerated the incommensurability between successive paradigms, at least when applied to science after the scientific revolution. Kuhn did not differentiate between the case where an old theory is
believed to be strictly speaking false but still respected as a useful approximation and when a theory is abandoned as wrong tout court. For example, Newton’s theory is still considered good science even though we now hold it strictly speaking not correct, as a contrast to Aristotle’s mechanics, which is false and also considered to be useless in scientific practice. The change from Aristotelian to Newtonian mechanics is thus, in a certain sense, much more profound, than the change from Newtonian mechanics to relativity theory and, in the field of micro physics, the advent of quantum theory. Furthermore, I do not believe that scientists working in different fields or research programmes have completely different norms; on the contrary, they share several norms.

In (Dawid 2009), it is argued that the conflicting assessments of string theory is due to a difference in Kuhnian paradigms. I find this to be an exaggeration. I do not believe that there is a radical difference in methodology or criteria for evaluating success between defenders of different research programmes in quantum gravity and traditional physicists. All involved agree about the need for empirical testing and there is no sign of there being radically different interpretations of empirical observations, which was a characteristic part of Kuhn’s argument for there being incommensurable paradigms. Dawid talks about a ‘metaparadigmatic shift’ since even the methods have changed and not just the view of the world. This talk of ‘metaparadigmatic’ is not needed since methodological norms are included in the concept of a paradigm and Kuhn held that methodological norms could differ between different paradigms.

4.3.1 Why are tests of string theory required?

String theory has not, so far, predicted any new observable phenomenon and it is understandable that people lose patience. On the one hand, this criticism is serious; a theory that can’t predict anything new appears to be no better than pseudoscience. But on the other hand, string theory’s basic aim is not just to make novel predictions but to explain hitherto unexplained facts, so why would lack of testable predictions be so problematic? Has the critics misunderstood the goal of string theory?

I think not. More than consistency and explanatory power is needed as criteria for selecting the correct explanation, since explanatory power is much too vague; there is no agreement about what an explanation is, not even how to compare two explanations and telling which is the better one.

Most theoretical physicists do not seem deeply concerned about the present lack of testability. Of course, they all would like a novel testable prediction coming out true, but the lack of testable consequences have not, at least so far, convinced the majority of those doing string theory to give up. Why not? I don’ think that they have dismissed testability as a necessary criterion for
empirical science, quite the contrary. String theorists simply hope to be able in the future to derive testable consequences.15

Some critics might say that instead of looking for reasons given by the researchers themselves we had better look for external factors that in fact cause them to continue work within it, causes that do not provide scientific reasons. Looking for reasons is to adopt an internalist perspective, asking for causes that are not reasons is to adopt the externalist perspective.

4.3.2 String theory from an internalist perspective: explanations
The three prime goals for pursuing string theory are, to repeat, i) to unify all four forces of nature, ii) to explain the vast particle diversity, iii) to derive or at least explain the constants of nature from first principles. Fulfillment of all these goals, not only the second one, could plausibly be called explanations, so the reasons for doing string theory can be formulated as quests for explanations. This motivates a look at what kind of explanation is suitable in fundamental physics.

It is clear that the hope for an explanation embracing all four fundamental interactions (a ‘Theory Of Everything’) is strong impetus for those doing research in string theory. But how does a satisfactory explanation look like? The answer depends on ones metaphysical position; if your inclination is pragmatist or instrumentalist, holding that all we want of science are reliable predictions, the difference between explanation and prediction is only a difference concerning timing. We predict what will happen and explain what has happened, but this is only a pragmatic difference; the logical structure is the same, often held to be captured by the Deductive-Nomological model of explanation, or its statistical counterpart.16

But if you hold that all scientific propositions are descriptions of the world that are true or false, then explanations, in some sense stronger than mere true predictions/retrodictions, are eagerly wanted. Most people are realists in this minimal sense and so for example they might ask themselves: is the correct explanation of particle diversity the one in terms of strings, i.e., one-dimensional physical objects?

Explanation as unification
A possible model of the explanatory structure of string theory, and of any fundamental theory, is unification. There are at least two explications of the notion of unification, (Friedman 1974) and (Kitcher 1981). However neither of these are entirely satisfactory. Albeit it’s shortcomings, I hold Friedman’s

---

15 For an example of a text where a physicist discusses the need for string theory to confront empirical tests see (Schnitzer 2003).
16 An explanation is, according to the DN-model a derivation of the explanandum from laws and other assumptions. In the statistical version the laws are statistical.
attempt to be the better one. It’s general idea is easy to grasp: a statement, or a set of statements, are explanatory and functions as explanans if it (they) enable derivations of several laws and empirical regularities that previously seemed unrelated and already accepted as being true.

It seems plausible that the prospect for such unificatory power is the basic reason why many theoretical physicists are so fond of string theory. The basic idea of string theory, that the fundamental objects are one- or multidimensional objects whose internal vibrations explain their observable properties carries great promises for unification.

The first aspect of string theory’s unificatory power is its ability to bring in gravitation into a quantum theory of the world. This tells very much in its favour, in particular since it was unintended. This reasoning has the form of a ‘theoretical no miracle argument’: it would be a miracle that string theory is able to unify the standard model and gravity if it is not substantial truth in it.\textsuperscript{17} In contrast to the postulate of extended fundamental objects however, alternative ways of unifying gravitation and the standard model is not a priori excluded.

Since GTR was known in advance, one cannot literally say that string theory predicts it; it is rather a case of retrodiction. But this is precisely a consequence of having proposed a unifying assumption.

The situation is analogous to that of GTR at its initial stage: GTR entails that the perihelion of Mercury will precess. This precession had been observed since long but no satisfactory explanation was available before GTR. Since GTR was constructed from a few basic principles, its ability to explain a known but hitherto unexplained fact count strongly for its at least approximate truth.

**Explanation of diversity of objects**

Lee Smolin, albeit a critic of string theory, frankly acknowledges that string theory explains the diversity of matter and force particles. The explanation has the form of unification, in a somewhat vague sense. In short, assuming that everything is made up of strings that propagate according to one simple law is all we need to get started. He writes:

\begin{quote}
Indeed, the whole set of equations describing the propagation and interactions of the forces and particles has been derived from the simple condition that a string propagates so as to take up the least area in spacetime. The beautiful simplicity of this is what excited us originally and what has kept many people so excited; a single kind of entity satisfying a simple law.

(Smolin 2006b, p. 184)
\end{quote}

It is easy to feel the pull of such a forceful explanation, and I think Smolin is right in saying that this beautiful simplicity, together with its using previously

\textsuperscript{17}Compare this with the ordinary no miracles argument, in support for scientific realism, which will be discussed in the following chapter.
well-known physical principles, is what makes people continue to hold on to
the string theory programme.

The epistemological structure of this explanation is as follows. It is not the
case that we knew in advance that there are strings that propagate according
to a certain law, and then physicists were able to show that many phenom-
ena thus could be explained. The epistemic order is rather this: we have a
problem, a confusing diversity of different kinds of objects, each kind fol-
lowing specific interaction laws. Then someone suggests two hypotheses: i)
all objects are strings, ii) these objects propagate so that the covered area is
minimized. Then, using the usual method of quantization, we can derive the
formerly unrelated phenomena. These derivations provide good reason to ac-
cept the two hypotheses as true. If so, we have an explanation in the form of
unification.

The strength of the explanation appears to be strongly related to how diver-
sified the things explained are; the more seemingly unrelated phenomena that
can be derived from these two assumptions, the more probable these assump-
tions appear to be, and the better unification we have.

**Explanation of constants**

From the basic theories in physics, the standard model and GTR, one cannot
derive the value of the constants of nature; rather, their values seem to be
completely accidental, and very lucky ones, because only a minute change in
any of them would make it impossible for any form of life known to us. In
short, it seems to be a miracle that the physical conditions for life in our world
are fulfilled. From this perspective a new and demanding explanation request
come to the fore: how do we explain the values of those constants?

Since the original hope of being able to derive the unique and correct values
of these constants from first principles had to be abandoned, the present idea
is to say that different combinations of parameter values exist and correspond
to different universes.

That means that our world is just one of staggering number of worlds, the
number $10^{500}$ has been mentioned, (Susskind 2005, p. 290). These worlds are
not supposed to be just mere possibilia, but as really existing! Of all these
existing worlds, we live in one, and this is explained by the fact that of all
possible combinations of values of constants, only very few permits life as we
know it. This is the anthropic argument. Many physicists feel quite uneasy
about it; is this really a viable explanation?

My view is that it’s validity rests on holding those worlds as really existing;
if they were mere possibilities, the argument would be a teleological one. For
suppose we take all the worlds as mere possibilia. Then the first step in the
answer to the question, “why do the constants have precisely the values they
have?”, would be that only our world is actualized. Why is that so? This is ex-
plained by saying that this is so because we live in it. This form of explanation
would do if we believed that our existence is the result of purposeful actions.
But most physicists don’t accept such an argument as a physical explanation and I agree with that assessment.

On the other hand, if it is believed that many of these different worlds really are existing, the explanatory structure is different. In this perspective no purposeful actions are assumed and different worlds with different possible combinations are realized. Then life evolves, as a matter of course without any intervention, in those worlds where conditions are suitable, and our world happens to be one such. In short, the anthropic argument makes physical sense only if it is taken for granted that many possible worlds are really existing. However, one could then argue that assuming the real existence of all these worlds reduces the explanatory value because now we have more assumptions in explanans.\footnote{18}

Summarizing, string theory brings about unification all four forces of nature and of the diversity of particles. But the third explanatory demand, viz., explaining the constants of nature is more problematic.

Other applications of the mathematics of string theory
In addition to the demand for unification of fundamental physics, research in string theory has also been justified by its contribution to the development of pure mathematics, in the sense that it suggests interesting conjectures fit for being proved more rigorously. These things are relevant for justifying continued work on string theory, viewed as \textit{uninterpreted mathematics}, but provide no relevant argument for or against the correctness of string theory as a unified theory of all the fundamental interactions.

A similar remark can be made concerning attempted applications of the formalism of string theory to other topics than fundamental strings and quantum gravity. For example, using the mathematical results behind the understanding of dualities, attempts have been made to apply the formalism of string theory to other things. Dualities are seemingly different descriptions and formalisms that have physically equivalent content. For instance a theory of strings can be physically equivalent to a theory of particles and a type of string theory on one background can be physically equivalent to a different kind of string theory on another background and so forth. Exactly what this means and how this should best be interpreted is a complicated and interesting question that will not be discussed in this chapter.\footnote{19}

One example is the use of string formalism for describing quark-gluon plasma; see for instance the review (Gubser 2009) and references therein. If such attempts would be empirically successful, it would show the usefulness of the \textit{mathematics} of string theory; but it would not say anything about the string theory interpreted as a theory about all fundamental interactions or its ability to solve the three fundamental problems stated in 3.2.

\footnote{18}{The topic addressed here is discussed in much more detail in chapter 8.}
\footnote{19}{It will instead be discussed in chapter 6.}
4.3.3 String theory from an externalist perspective

Now let’s have a look at string theory from a perspective in which scientific activities are not viewed as purposeful deliberations and actions but as effects of causes described in a non-intentional idiom.

In the debate concerning string theory, some critics have suggested that non-scientific factors must be taken into account when trying to understand why people have not left the business, despite its lack of empirical success. Assuming that empirical testability is a *sine qua non* for empirical science, there must be something else, i.e., factors external to the science per se, that has made people willing to pursue this research.

Another argument put forth in, for instance (Smolin 2006b), is that now almost all important positions in theoretical physics at major universities are held by string theorists. Students are forced into string theory due to lack of alternatives. It is claimed that other approaches to deal with the problems of quantum gravity is not given enough funding and that this is not fair. Is this a telling explanation?

As is well known, string theory was for quite some time pursued only by a few number of individuals. They had little funding, being far beside the focus of interest in theoretical physics. But they came up with results that interested a wider audience and within a short time string theory became mainstream. Today other approaches to quantum gravity have no less funding than string theory had at the beginning. There is reason to think that if people in a programme competing with string theory would come up with good enough results, their approach would most plausibly attract more researchers and get better funding.

The crucial question is, of course, who is to judge what to count as a ‘good result’; do adherents to different approaches share the same, or sufficiently similar, norms regarding what to count as progress, or do they differ? If they have sufficiently similar norms there is hope for a fair game. I think that that is in fact the case.

To make the argument by Smolin more convincing one needs to make it plausible that string theorists of today are more stubborn than the previous generation of particle physicists who turned from their previous work into string theory. One possible argument for that can be put like this. The work done in particle physics using the framework of QFT produced many testable results and led to many successes. When scientists previously working in this field turned to string theory they did not thereby nullify the worth of their earlier work. In contrast, if a string theorist would give up and move to another area of research he would thereby publicly show that he thought string theory was a failure. Such a move is plausibly more difficult to take if ones whole career has been confined to string theory.

However, as has been mentioned above, I do not believe that string theorists have abandoned the idea that experiment is the final arbiter of theories; there is
still in the physics community a much higher valuation of empirically testable and confirmed predictions than merely theoretical results, however interesting.

4.4 Other approaches to quantum gravity

Here I add a few words on other approaches to quantum gravity. Much of what has been said concerning string theory applies, with few modifications, to all approaches to quantum gravity. Everything, that is based on the lack of empirical progress, can be applied without change. There are however a few points made concerning string theory that do not apply to the other approaches to quantum gravity.

One such issue is unification. The other approaches do not share the ambitious goal of string theory to completely unify all forces. They are happy to formulate a quantum version of gravity that could stand alone without considering the other forces. To a smaller extent this is also a kind of unification since they want to describe gravity in terms of quantum physics. Nonetheless this is not the same as unification in the stronger sense. It is more a question of compatibility between theories.

Another issue is how strong or dominant the approaches are. Here it is, as has been mentioned before, clear that string theory has the upper hand. It has by far the largest research community. Other approaches have much fewer followers, sometimes they are only worked on by one or two researchers.

4.5 Conclusions

In this chapter I have discussed the present status of string theory, and to a much lesser extent other approaches to quantum gravity, from different methodological views, but also elucidated these methodological views in the light of an ongoing research programme, about which it is not known know whether it will be a success story or a dead end.

The long debate in general methodology has not produced anything like consensus concerning norms for rational choice between competing theories, paradigms or research programmes in the same field. This negative outcome is in my view to be expected and to be welcomed. This is so because if there were consensus about a set of universal and permanent methodological norms telling us when to keep on and when to reject a research programme, these could only have been motivated by a priori arguments and independent of actual varying practices. That would mean that philosophers of science would have agreed upon a kind of first philosophy, an epistemology independent of empirical and thus revisable considerations.
I believe that such a position is untenable. I wholeheartedly accept that epistemology is part of empirical science and is open for revision, as all empirical science.

It is common to argue for a particular methodology by using success stories from history of science, which by their historical nature are such that we know the outcome and we know that the scientists made the ‘right’ decisions. It has, for example, since long been agreed that the atomic theory of matter is true and those who opted for it took the correct decision; they were rational. So one can raise support for a particular methodology by describing this theory such that it fits the general methodology. But the force of the example comes from the success. Looking at an ongoing research programme/paradigm/theory such as string theory whose fate is unknown, shows that none of the well-known views in general methodology gives a clear recommendation on how to continue.

I reject the idea that philosophers of science can act as jury members in telling scientists whether a research program should be given up or not. Philosophers have no such privileged stance from which to decide such things. The role of philosophers of science is to be active participants in an ongoing discussion on science and scientific methods but not to lay down the law once and for all.

There is little reason to claim that theoretical physics in general, and string theory in particular, has lost its aim to predict outcomes of new experiments. String theorists would of course consider the derivation of a testable prediction as big success and if the prediction would come out true, the victory would be telling. That string theory so far has not been able to come up with any testable consequences is disappointing, but not reason to lessen the demands on physical theories and string theorists themselves show little sign of changing their mind on this issue.

See (Hedrich 2006) for a somewhat different view; he argues that with the introduction of the ‘Landscape’, research in string theory has become more like metaphysics than physics, and he suggests that string theorists have been less concerned with the lack of empirically testable consequences in recent years. I agree that some string theorists are very speculative, but it does not seem correct to say that string theorists in general have given up the hope of testability. But of course, the longer it takes for string theorists to come up with anything testable, the more speculative the endeavor appears.

The recent suggestions that there really exists a multitude of universes, the ‘Landscape’, seems wildly speculative. The critic might reasonably say that the only reason to believe in the real existence of a multitude of universes is that it explains, in a certain sense, the values of the constants in our universe. But this explanation only increases my wonder; in what lies the explanatory value of postulating a multitude of universes, enormously many more than the number of constants to be explained? It’s hard to see any simplification or unification in this; what is needed are some independent arguments for the reality of the multiverse. It should also be noted that for there to be any kind
of explanation it must be assumed that the other universes are actual and not mere possibilia. The idea of the ‘Landscape’ of universes has also been controversial within the physics community.  

A philosophically interesting conclusion to be drawn from the fact that string theory still dominates theoretical physics is that those involved in general have rather strong realist convictions. If they had no such convictions their interest and stubbornness in pursuing this research programme would be inexplicable.

Popular accounts of string theory, of which a large number has been published in recent years, have not always been sufficiently clear about its speculative character. It is important to stress, that string theory is not empirically well tested and part of established science, in the way that GTR and the standard model are.

As already indicated, I think Lakatos theory, MSRP, in Hacking’s interpretation, is the most viable view in methodology. But there is one aspect lacking in Lakatos account, viz., the possibility of a merger of different research programmes. There has been some such examples in the history of science. One example is Schrödinger’s wave mechanics and Heisenberg’s matrix mechanics. Also string theory itself is an attempt to merge earlier research programmes. The possibility of a similar merger of ideas from string theory and from some of its present competitors such as loop quantum gravity should be kept in mind. Different research programmes might be holding different parts of the puzzle. This I consider to be a good reason for arguing for a more open attitude between proponents of different research programmes especially when no programme is truly progressive at the empirical level. Basically all programmes are equally bad when it comes to empirical progress.

The situation for those who decide about funding of research in theoretical physics is somewhat similar to investors in venture capital companies who want to invest money in projects that will be successful in the future and avoid waisting money on bad ideas. But no one knows in advance what will be successful and there are good reason to think that most projects will fail and few will succeed. So the venture capitalist had better not put all eggs in the same basket, but to engage in several projects, in the hope that one or two of them will succeed and the profits in these will cover the lost money in the others.

20See chapter 8. for more details.
21The view that it was just found out that two formulations of quantum mechanics were equivalent, is severely simplified. This has been argued in detail in (Muller 1997).
22See also the discussion on merging in (Audretsch 1981), where he discusses quantum gravity in general.
23To avoid any misunderstanding I would like to state the following. The goals of the venture capitalists and of the research funders are assumed to be different. The goal, in the case with funding of science, is not supposed to be monetary gain but empirically successful theories.
The parallel seems sufficiently close for a similar conclusion regarding funding of research; we had better not use all money on one idea, but to diversify. There are no strong reasons to think that string theory is the only way forward; another approach might very well be the correct one, and it cannot be known in advance which.

From my perspective I believe that a more pluralist approach is important in a situation where no research programme is empirically progressive. Still I think it is to be expected and even recommended that if one research programme starts to be empirically progressive this programme will receive most funds until it is again stuck and does not produce new empirical results. This I think is in general a reasonable strategy to use for people providing funds.

For a new programme to become progressive there need to be risk takers that choose to work with a programme with dim prospects. To start a new research programme or early on join a new research programme is a high risk strategy. You will be hailed as a genius if you succeed and forgotten or seen as a crackpot if you fail, just as the risk taking entrepreneur that tries something new might become very rich or fail miserably. In both cases predictions are highly uncertain.

Just as it would not be good for the economy if we all were risk taking entrepreneurs it would not be good for the scientific community if all scientists were iconoclastic revolutionaries. A mix of risk takers, critics, and routine workers is needed. The scientific community needs different kinds of scientists working on different approaches and with different attitudes. Feynman expressed an analogous view some decades ago in his Nobel lecture:

Therefore, I think equation guessing might be the best method to proceed to obtain the laws for the part of physics which is presently unknown. Yet, when I was much younger, I tried this equation guessing and I have seen many students try this, but it is very easy to go off in wildly incorrect and impossible directions. I think the problem is not to find the best or most efficient method to proceed to a discovery, but to find any method at all. Physical reasoning does help some people to generate suggestions as to how the unknown may be related to the known. Theories of the known, which are described by different physical ideas may be equivalent in all their predictions and are hence scientifically indistinguishable. However, they are not psychologically identical when trying to move from that base into the unknown. For different views suggest different kinds of modifications which might be made and hence are not equivalent in the hypotheses one generates from them in ones attempt to understand what is not yet understood. I, therefore, think that a good theoretical physicist today might find it useful to have a wide range of physical viewpoints and mathematical expressions of the same theory (for example, of quantum electrodynamics) available to him. This may be asking too much of one man. Then new students should as a class have this.

That these might, sometimes at least, be used in economically profitable ways, is not something that I care about here.
If every individual student follows the same current fashion in expressing and thinking about electrodynamics or field theory, then the variety of hypotheses being generated to understand strong interactions, say, is limited. Perhaps rightly so, for possibly the chance is high that the truth lies in the fashionable direction. But, on the off-chance that it is in another direction - a direction obvious from an unfashionable view of field theory - who will find it? Only someone who has sacrificed himself by teaching himself quantum electrodynamics from a peculiar and unusual point of view; one that he may have to invent for himself. I say sacrificed himself because he most likely will get nothing from it, because the truth may lie in another direction, perhaps even the fashionable one.

(Feynman 1972 [1965]b)
5. Realism, semantics and scientific theories

The purpose of this chapter is, primarily, to make the two following chapters easier to read for people who lack the needed background in philosophy of science. To some extent this will lead to some repetitions in following chapters. This chapter deals with issues concerning scientific realism and semantics of theories. To some extent the discussion on realism has been addressed also in the previous chapter. While this chapter is mainly meant as background for the following discussion I will already here express some of my own views on scientific theories.

5.1 Introduction

After the fall of logical positivism, new views on the nature of scientific theories, have been developed within philosophy of science. The logical positivists tied their views on scientific theories closely to semantics. In various ways, the idea was that for an expression to be meaningful or cognitively significant it had to somehow connect to what was empirically accessible.\(^1\)

In many newer views on scientific theories, questions concerning semantics do not play such a significant role. Scientific realists and constructive empiricists alike, share the following view. Theoretical statements, if precise well formed and apparently understandable, can be judged cognitively significant even if they could not be connected with experiment. They say that scientific theories are to be understood ‘literally’. But what does that really mean? I think, that in these new ways of looking at scientific theories, questions of a semantic nature has been treated too carelessly and I think that questions concerning semantics must be readdressed in philosophy of science.

I do not want to say that philosophers have not cared at all about semantic issues regarding science. It is just that many of the contemporary views on scientific theories, especially standard accounts of scientific realism as expressed in (Psillos 1999) and constructive empiricism as it is expressed in (van Fraassen 1980), treat semantic questions very briefly and basically just assume that it is OK to understand statements in physical theories in a straightforward and literal way.

I will start by giving a brief overview of previous positions so as to put the discussion in context.

\(^1\)See for instance (Suppe 1977), (Ayer 1952) and the collection of original articles in (Ayer 1959).
5.2 Earlier views on scientific theories

Previous views, on scientific theories, have had different attitudes towards questions regarding semantics and realism. In this section I will present a few different views on these issues.

The question on realism, regarding scientific theories, is about to what extent we ought to believe what our most successful theories tell us about the world and its constituents. Should we believe in the existence of the different entities or objects that are posited in scientific theories even though they are not directly observable? Here I consider things described with terms such as: ‘electron’, ‘quark’, ‘string’, ‘quantum field’ and ‘wave function’ do they refer to something in reality? One need not come to the same conclusion regarding all these terms. It is possible to consider arguments for the existence of some type of entity to be more convincing than the arguments for another type of entity. Debates on these questions have appeared many times during the history of science, sometimes a consensus has appeared that we should believe that certain entities exist. Think of the debate regarding atoms. Other times a theoretical substance or entity was taken seriously for quite some time but was later disregarded as a mistake, examples are flogiston and the luminiferous ether.

Note however that even though a consensus was formed, some might deem this not to be sufficient. Anti-realists find only the truly empirical content of a theory to be important and consider the real value of a theory to be only its empirically accessible predictions.

5.2.1 Logical positivism and verificationist semantics

To some extent, I here repeat things that has been told about the logical positivists in the previous chapter. The focus here is however different and I find this repetition to be useful for stressing the points I want to make concerning semantics that are developed here and in the two following chapters.

Logical positivists are famous for their verification theory of meaning. One could say that they endorsed a verificationist semantics. Due to this their views on science, epistemology and semantics are intimately tied to each other. The distinction between analytic and synthetic statements was very central to their scheme.

The logical positivists did not care about ontology and found speculations regarding metaphysical issues to be basically meaningless. What was really there behind what we could measure, was considered to be irrelevant and meaningless. Talk of atoms, for instance, was a convenient way of formulating a theory for making predictions. No commitment to the existence of atoms was needed.

Logical positivism is however not a popular doctrine among present day philosophers of science. Their verification criterion for meaning has definitely
fallen in disrepute. The story concerning the rise and fall of logical positivism has been told by many before and I will not repeat it here. For an extensive description of the formulation, development and problems that faced logical positivism I refer to (Suppe 1977).

Their ideal was an axiomatic formulation of theories and they also supported a syntactic view on theories. By this it is meant that they viewed theories as defined by a set of sentences that could be derived in an axiomatic formalism. Then they needed so called ‘bridge rules’ or ‘correspondence rules’ to make contact with empirical reality.

5.2.2 Quine and semantic holism
Quine also views theories in an axiomatic and is a supporter of the syntactic view on theories. There are however important differences with the logical positivists. He does not accepts the distinction between analytic and synthetic statements and denies that a sharp distinction, in language, can be made, between terms referring to the observable and the non observable.

Despite this his views on semantics are still tied to the empirical. However he does not think it possible to make this link directly on individual terms. It is the theory as a whole or at least ‘chunks’ of theory that is connected with the empirical. It could be said that Quine keeps a verificationalist semantics, but within a holistic context.

Quine accepts talk about ontology. He thinks that we make ontological commitments beyond what is directly observable when we hold a theory to be true. His main idea in this area is that ontological commitments are to be read off from formalized theories. Everyday language is deemed too imprecise to be used for deciding what kind of ontology to endorse. He claims that,

The common man's ontology is vague and untidy.
(Quine 1980, p. 9).

He also repeatedly states that “to be is to be the value of a (bound) variable”. For this to be clear the theory needs to be formalized in first order predicate logic. This will make it apparent what we quantify over. It is important to note that this is not supposed to be an argument for which objects that exist. The purpose is different, according to Quine it is only supposed to make clear what we are committed to when we accept a theory as true.

One problem, according to me, with Quine is that his view on theories is way too idealized. Real theories in science, even as mathematically formulated as they are in physics, are not described with the logical precision he assumes. Nor is it reasonable to believe that ideal to be attainable. However despite this I think there is quite a lot to be learned from Quine’s arguments. The cases, where I think his conclusions can be trusted, are when we get negative results for the idealized theories he is dealing with and where it also seems that if
something does not even work in the idealized theory, it is even less likely to work without the idealizations.

Quine understands ontology only in terms of what objects a theory quantifies over. But his view on objects is very theoretical, he does not view objects to be independent of the descriptions they get in theories. He also allows for reformulations of theories using proxyfunctions. A proxyfunction maps objects in one theory formulation to objects in another, while keeping the overall structure of the theory the same. This has no deep importance for Quine. The real referent of a term is not that relevant. We can not really distinguish them and if they are different, they are so in an insignificant way. He writes,

What makes sense is to say not what the objects of a theory are, absolutely speaking, but how one theory of objects is interpretable or reinterpretable in another.
(Quine 1969, p. 50)

Structure is what matters to a theory, and not the choice of its objects.
(Quine 1981, p. 20).

Save the structure and you save all.
(Quine 2008, p. 405), in which (Quine 1992) is reprinted.

All these statements seem to go in the direction of defending some form of understanding of theories based on structures. This is similar to the structural realism that will be discussed in the following chapter. However he seems firmly wedded to a view where structures themselves in no way can replace the objects as the basis for understanding ontology. He writes for instance,

My global structuralism should not therefore, be seen as a structuralist ontology. To see it thus would be to rise above naturalism and revert to the sin of transcendental metaphysics.
(Quine 2008, p. 406), which (Quine 1992) is reprinted.

This is a “sin” that some structural realists, whose views will be described in the next chapter, would happily make. Also, I can not myself understand what Quine finds so problematic with a structuralist ontology, given all the above quoted remarks.

5.2.3 Arguments for scientific realism
Scientific realists claim that we have good reasons to believe in the existence of the entities posited by successful scientific theories. By successful it is here understood that they, at least, are empirically adequate and hence can predict experimental results correctly.
It is also typical for scientific realists to argue that theories should be interpreted literally, they see no problem with semantic questions regarding highly theoretical discourse in science.

One very common argument for realism is the so called no-miracles argument. This is most often attributed to Putnam, see for instance (Putnam 1979, p. 73). In short this argument states that it would be a miracle if our theories were so successful in making new predictions if they did not at least in part say something about what really exists in the world.

One can discuss the question on how important it is that the predictions are new. It has been argued that it can be sufficient that the theory find new explanations to experimental results, even though they were known before. This would be acceptable as long as the theorist did not intentionally design the theory to explain these experimental results. A classic example, mentioned also in the previous chapter, is how the general theory of relativity was used to explain the observed orbit of Mercury. It was already known that the orbit did not behave as expected according to Newton’s theory of gravitation before the explanation was given in GTR.

Another argument in favor of scientific realism was given by Ian Hacking. The argument focus on the experimental side of research. The point of this argument is that, when we have technology that can do what would seem like direct manipulation of the theoretically posited entities, it seems absurd to deny that they exist. In these cases we can trust that the entities described in the theories exist, even if the theories themselves may be wrong in other ways. An important point is that; the entities that we can manipulate are not defined in terms of the theories. Hence their existence or nonexistence does not depend on the constancy of our theories or the exact theoretical description we give of these entities. The point that different theories can refer to the same entities, even though the theoretical description of these entities differ has also been defended by Putnam, his argument are however not clearly related to the importance of experimental work.

Further, more detailed, arguments for different versions of scientific realism can be found in (Churchland & Hooker 1985).

5.2.4 Arguments for scientific anti-realism

Van Fraassen advocates a modern anti-realist view on science called “constructive empiricism”. It is important to understand that van Fraassen is a semantic realist; he thinks that statements in a theory can be true or false, regardless of our ability to decide which is the case. He does however not find that the aim of science is to find truth on all levels but only to find empirically adequate theories.

(Hacking 1983, p. )

See Putnam’s ‘The Meaning of “Meaning” ’ in (Putnam 1975).
What are the reasons for supporting an anti-realist view concerning scientific theories?

Nowadays most physicists consider the existence of atoms and electrons to be beyond doubt. This is due to immense experimental data, that seem to confirm the existence of these entities. But even when we have such impressive data and empirical success, there is still a possibility to question the existence of these entities. It could be said that we do not immediately see or experience these objects, what we see are rather effects that are consistent with the assumption that there are atoms and electrons.

Instrumentalists like the logical positivists, and other kinds of anti-realists, do not question the value of theories that posit electrons or atoms. They admit that these theories are scientifically respectable and useful tools or instruments for predicting empirical effects; they merely claim that we are not, by that fact, justified in inferring the existence of the posited entities. Or at least that it is unnecessary and irrelevant to believe in their actual existence.

An important argument for anti-realism has been based on underdetermination of theories by data. This argument points out, that we cannot rule out the possibility that many different theories can be consistent with data. To argue for the thesis of underdetermination, the first step is to remind ourselves of an elementary rule of logic. That is, we can not deduce $p$ from $p \rightarrow q$ and that it is observed that $q$ holds. The logical possibility remains that there might be an alternative explanation for the correctness of $q$. All this is of course more than well known and the logical mistake is famous under the name ‘affirming the consequent’. Even though this is well known to philosophers, and obvious to everyone who reflects upon it, it seems sometimes to be forgotten when science is discussed.

Often one finds the evidence so convincing since one cannot conceive of any alternative explanation. But the exact nature of a theory cannot be strictly speaking proven. Previously held theories have been replaced by new theories. It can then be claimed that these new theories, actually were possible alternative explanations to the old data all along. This brings us to another closely related argument for anti-realism, namely the pessimistic meta induction. History of science tells us that even very well confirmed theories, such as Newtonian mechanics, is sometimes overthrown. The new theory, which replaces the old, suggests a radically different view of the world than the previous one. Our conceptual framework and the view of the world have for instance been radically altered due to quantum physics and the theories of relativity. It is of course still the case, that we use and trust Newtonian mechanics to give correct empirical predictions, within very small margins of error in ordinary circumstances. Newtonian mechanics is still very usable as an instrument for (approximate) predictions; we can however no longer strictly speaking be said to believe in the world-view presented by Newtonian mechanics.

The point is that when two theories both are consistent with data, at a certain time, there is a form of effective underdetermination but given future develop-
ment of experimental techniques the theories might actually be distinguished. To show that two theory formulations could never be distinguished in the future one need to show how they are related. This is needed to show that they give exactly the same predictions regardless of the situation empirically accessible at the time or otherwise. In this case it could be argued that they are not really two theories but the very same theory formulated differently. The conclusions drawn are however tied to what view on semantics that are taken for granted. I will come back to these questions in the following chapter.

Old time positivists, who according to the modern terminology are seen as anti-realists, considered only the empirical content of a theory to be meaningful. They could handle underdetermination, since a difference at the theoretical level is not significant according to them, only differences in empirical predictions need to be considered. The pessimistic metainduction on the other hand has been used as an argument against the positivists assumptions that science is cumulative. But if one denies extreme forms of incommensurability and admits that there is a possibility to compare empirical predictions, even if they come from different conceptual frameworks, this argument can be countered. In practice then a logical positivist can treat the underdetermination and pessimistic metainduction in a similar fashion. If you do not really care about all the theoretical stuff, in a theory, it is really not such a big deal if the theoretical framework is changed.

5.3 A literal understanding of theories?

Many modern views on science argue that we should understand theories literally also when they describe things that are far removed from what we directly observe. A problem here is how we are to understand the statement that theories are supposed to be interpreted literally. What does that really mean? To what extent can we understand theories literally when they attempt to refer to things so far from our ordinary experiences? In the next chapter this question will be readdressed, there a more explicit description between the difference between semantic and epistemic realism will be given.

One part of claiming that a theory should be understood literally is that all statements, regardless of whether they refer to what is observable should be thought of as having well defined truth-values.

It could also be said that a technical terms, such as ‘electron’, must be understood literally in one sense, since a new word such as ‘electron’ does not have another older meaning. It is hence not a metaphorical use of the word ‘electron’ that is used. This is however a side issue and is not really relevant for the point I would like to make, concerning a ‘literal’ understanding of theories.

---

4For arguments of this kind see for instance (Quine 1975) and (Dawid 2007).
What I find relevant, is rather that when we try to explain what a technical term is supposed to refer to, we generally use analogies to more everyday experiences. This is done to convey some sort of description of these objects. These descriptions should certainly be taken with more than a grain of salt, hence not literally. A more precise explication of the notion can be considered to be implicit in the mathematical formulation, and its connection to experiments, of the theories hence the meaning of a term is decided by its use in a more developed framework.

Nevertheless, the pictures are not unimportant for our ability to gain an understanding of the theories, they are often important heuristic tools. It is also quite difficult, in real theories of physics, to read directly from the formalism what parts of the theories that should be interpreted to refer to something physical and what is a purely mathematical superstructure. This issue will be investigated, using string theory as an example, in the following chapter.

5.4 Final comments

Since there are strong arguments, on both sides of the realism vs. anti-realism debate, in philosophy of science, it seems that some form of intermediate position would be the best. In the following chapter the discussion concerning scientific realism will be continued, with some minor repetitions. There the discussion will be connected to the dualities found in string theory. I will also discuss various species of so-called structural realism. The different kinds of structural realism would all be considered to express such intermediate positions in the debate on scientific realism.

Based on these, I will argue for a milder form of verificationist semantics than the one espoused by the logical positivists. It also differs from the kind of verificationism, connected with semantic holism, that can be found in Quine’s writings.
6. Realism, underdetermination and string theory dualities

This chapter is directly based on (Matsubara forthcoming). Except, for the elimination of some superfluous background material on string theory, only very minor modifications to that text has been made.

6.1 Introduction

The main topic of this chapter is the interpretation of the so-called ‘dualities’, that play a prominent role in string theory. For now it suffices to say that we have a duality when two described systems that prima facie seem to be very different still are physically equivalent; or maybe even not two systems at all but rather just one system under different descriptions. Various ways of understanding dualities will be presented in this article and different interpretations are given depending on various views on semantic and epistemic questions.

For those who take string theory seriously the work is of relevance to their understanding and interpretation of what string theory tells us about reality. The discussion on how to understand dualities is also of interest as an example of how to think about theoretical claims, it gives relevant input to the debate on scientific realism.¹

6.2 Underdetermination and scientific realism

Here I continue the discussion on scientific realism that was initiated in the previous chapter. Now the discuss will be much more closely tied to the discussion on string theory.

6.2.1 Underdetermination

The problem of underdetermination is that we cannot rule out that more than one theory is compatible with our empirical data. There are different kinds

¹Works closely related in topic to what is discussed in this text are (Dawid 2006, Dawid 2007) and (Rickles 2011). More on quantum gravity written by philosophers can be found in (Callender & Huggett 2001), (Rickles et al. 2006) and (Rickles 2008b), see also the comments in section 3.7.2 of (Ladyman & Ross 2007).
of underdetermination. We can talk about underdetermination with respect to currently available data; this kind of underdetermination is called ‘transient underdetermination’ or ‘scientific underdetermination’. Theories might differ in their predictions concerning what have not yet been empirically tested and still be underdetermined in this sense.\(^2\)

The other kind of underdetermination is between theories or theory formulations with respect to all possible data. This means that all their predictions are exactly the same. This is the kind of underdetermination that will be considered in this paper. How one responds to this problem depends to a certain extent on ones views on scientific theories. If one supports something like an instrumentalist position, and individuates theories only in terms of their empirical content, the problem disappears. If so one must consider the differences that seem to exist between various formulations to be without any real significance; the theories say the same thing. Our talk about the unobservable is just empty words, which should not be taken seriously. If, on the other hand, we assume that alternative theory formulations describe different scenarios, the threat of underdetermination must be taken as real.

Quine is associated with underdetermination and the claim that two logically incompatible theories can both be consistent with data. However, as can be seen in (Quine 1975), his views are quite complex. There he states that if there exists a mapping between two theory formulations, they do not describe different theories at all; instead they are to be understood as different variants of one and the same theory; Quine calls this ‘reconstrual of predicates’. So, he did not consider any two formulations that give rise to the same empirical content as being a genuine example of underdetermination. Note, however, that this does not mean that he rules out examples of genuine underdetermination that cannot be understood in terms of a reconstrual of predicates, but it can be difficult to find such examples.

Quine’s view, that theory formulations that can be mapped to each other using a reconstrual of predicates should be understood as different formulations of the same theory, would not be satisfying for all. It can be claimed that the formulations present two genuine alternative theories after all, despite the structural similarity. A person claiming this would give more importance to what is stated in the two formulations beyond what is captured in the structural properties of the formalism. They would find Quine’s views still too positivistic in spirit. They could argue that there are relevant semantic differences that are lost in the mapping; the mapping can only be done for that part of the theory that is logically or mathematically formalized. So here we note that we must ask if there is more to a theory formulation than what is captured in the logico-mathematical structure. This is a central question for this chapter.

\(^2\)(Sklar 1975), (Stanford 2001) and (Dawid 2006, Dawid 2007). Dawid argues that string theory and the dualities can be used as an argument against the importance of this kind of underdetermination. I do not address this question in this chapter.
6.2.2 Scientific theories and realism

Do our theories correctly describe the world even beyond what we can empirically measure or are they just tools for prediction? This has been debated for a long time and questions concerning realism have been discussed many times in the history of science and philosophy.

Roughly a scientific realist thinks that our best scientific theories are approximately true and that theoretical terms that are introduced in the theory typically refer to entities that really exists.³ Hence, we are justified in believing that there really are electrons. A scientific realist normally thinks that we should take what scientific theories say literally. Exactly how one should understand what we mean by ‘literally’ is not completely clear. Here it suffices to say that when using a literal understanding of theoretical statements it is assumed that they have semantically relevant content that goes beyond the empirical content. In short one can claim that there are two main components in scientific realism. Following Bain(draft) these are:

1. The semantic component: The theoretical claims of certain theories are to be interpreted literally.
2. The epistemic component: There are good reasons to believe the theoretical claims of certain theories.

To be a semantic realist is to accept the first claim and to be an epistemic realist is to accept the second claim. To be a traditional scientific realist you must accept both claims.⁴

It should be noted that a scientific realist would only consider a mature well-tested theory that has been used for novel prediction to be considered among those to be taken realistically. Obviously string theory does not yet, and perhaps may never, live up to this. So why would considerations of string theory be relevant for questions concerning scientific realism?

The answer is that the investigation in this chapter is concerned with what string theory would imply if it is taken seriously as a description of the real world. The dualities in string theory provide a very interesting case study where it is difficult to defend a traditional form of scientific realism. I also believe that the discussion elucidates questions concerning how modern theoretical physicists can interpret theoretical claims. I think this will be of value even if string theory will not live up to its expectations.

Underdetermination of theories by data is a problem for scientific realists. To solve this problem realists may argue that we can give good reasons for choosing one theory instead of another using virtues such as simplicity, lack of ad-hocness, explanatory power, etc. These criteria have been criticized as being vague and also nonindicative of truth, but I will not review this debate. Alternatively one can argue that what seems to be an example of underdetermin-

³See (Boyd 1985) or (Psillos 1999) for standard accounts of scientific realism.
⁴For a careful discussion and defense of scientific realism see (Psillos 1999). Another good source for the discussion on scientific realism is (Churchland & Hooker 1985).
mination are just two ways of describing the same theory. When using this approach the scientific realist must be careful not go too far in the instrumentalist or positivist direction; this would effectively turn him into an anti-realist.

In general, arguments using underdetermination are supposed to force the scientific realist to abandon either semantic realism or epistemic realism. I quote from Bain(draft) where the general form of an underdetermination argument is described like this:

For any version of semantic (epistemic) realism, there are theories \( T \) and \( T' \) such that if we are semantic (epistemic) realists about \( T \) and \( T' \), we cannot be epistemic (semantic) realists about \( T \) and \( T' \).

Traditional logical positivists, are a kind of anti-realists. They regard the cognitively significant part of a theory to be restricted to its empirical content. Hence according to that view string theory is not very impressive. On the other hand if string theory in the future would be empirically successful, the dualities would not cause any problems. They would just be seen as semantically equivalent since only the empirical content would be thought of as relevant. However, logical positivism is by present day philosophers of science judged to be an unacceptable view on scientific theories and I agree with this assessment.\(^5\)

A more modern form of anti-realism is defended by van Fraassen. In contrast to the logical positivists he does not tie his anti-realism to a theory of meaning. Just like scientific realists he claims that theories ought to be taken literally. He considers a well formed sentence to be true or false regardless of whether or not it is epistemically possible to decide what the truthvalues are; he is hence a semantic realist. But what is important for a scientific theory, according to van Fraassen, is only that it is empirically adequate, i.e. that it correctly predicts empirical data and hence ‘saves the phenomena’. We do not really need to believe what a scientific theory claims concerning what lies behind the phenomena. This is beside the point, since the aim of science is just to find empirically adequate theories; see (van Fraassen 1980).

Structural realism is an attempt to find a position between realist and anti-realist views on scientific theories. The modern discussion concerning structural realism goes back to (Worrall 1989).\(^6\) It is an ongoing debate how the doctrine of structural realism is to be understood and what insights it really gives concerning questions of epistemology and metaphysics.\(^7\)

Structural realists argue that ‘structures’ are preserved even when we radically change our theories of the world; the clarification of what these structures

---

\(^5\)For an extensive description of the formulation, development and problems that faced logical positivism see (Suppe 1977).

\(^6\)It ought to be mentioned that similar ideas have been formulated earlier. Worall himself argue that Poincaré advocated a position that could be thought of as a form of structural realism. No attempt is however made to discuss the historical roots of structural realism any further.

\(^7\)See (Frigg & Votsis 2011), for an detailed discussion on various forms of structural realism.
are supposed to be is unfortunately not very clear, even though it seems that
the structures are supposed to be intimately connected to the mathematical or
logical elements of a theory formulation. For example, equations in older the-
ories can be retained or at least be shown to follow approximately from new
theories.

It is possible to explicate the position in various ways. Ladyman (1998)
introduced the distinction between epistemic structural realism (ESR) and on-
tic or metaphysical structural realism (OSR). A proponent of ESR argues for
structural realism due to epistemic reasons; we can only know or have good
reasons for believing in the structural parts of a scientific theory. They do not,
in contrast to a proponent of OSR, think that we need to adopt a structural-
ist ontology i.e. assume that the structure as a whole and the relations in the
structure are ontologically more fundamental than the relata. So a defender
of ESR typically thinks that the structure supervenes on separate individuals
with certain properties and so forth. People have tried to explicate this notion
using Ramsey-sentences, but this approach is problematic.8

I will not assume that ESR must be understood in such a fashion that it
presupposes an ontology of separate individuals. This might deviate slightly
from the normal way of describing the difference between ESR and OSR.
A main difference, as I understand it, between ESR and OSR is whether one
considers talk about what is beyond the ‘structural’ to be meaningful in a more
substantial sense so that it can be used to describe real alternative situations.
In (Ladyman & Ross 2007), where OSR is defended, they argue for a form of
verificationism, but not a verificationism of the logical-positivists kind which
is tied very tightly to a semantic theory. I think that it is important to address
the issues concerning the semantics of scientific theories in more detail and
this will be done to some extent later on in this text.

In (Lyre 2011) it is asked whether structural underdetermination is possible.
Given that the notion of structure is not very clear it is of course difficult to
decide this. Lyre concludes that it could not be ruled out.

While I find structural realism compelling and am sympathetic towards
structural realism as a project, I find the lack of a reasonably precise and
useful description of structure in the context of scientific theories quite dis-
turbing. Due to this ‘structure’ might not be the best word to use. We could
use the expression ‘intermediate realism’ for any intermediate position in the
realism debate. Structural realism, if it could be explicated in a satisfying
manner, would be a species of this, but not necessarily the only possible kind
of intermediate-realism. Another kind of intermediate realism would be the
semi-realism which is developed in (Chakravartty 2007).

8This is discussed in (Psillos 1999) and (Ladyman & Ross 2007). See also the original texts:
(Newman 1928), (English 1973) and (Demoupoulos & Friedman 1985).
6.3 Formalisms and physical content

Mathematics plays an important role in formulating physical theories but mathematics by itself is not physics. Part of the mathematical formalism must be interpreted so that it refers to something physical. Details on exactly how this is supposed to be done will not be given. It should nonetheless be clear that I do not presuppose a traditional positivist line in which every concept must be given a definition, partial or complete, in terms of operations and measurements. So, for example, even though string theory has not been empirically successful, there is an understanding of how the theory is supposed to be related to earlier theories that have been empirically successful. For instance, one understands how the worldsheets of strings are supposed to be related to Feynman diagrams in quantum field theory and which calculations in string theory are supposed to be used to calculate scattering amplitudes and so forth. So string theory has some physical content. By ‘physical content’ I mean something less strict and less directly connected to observations and experiment compared to empirical content even though it at least indirectly should be connected to empirical content. I am aware that this distinction is vague but the example above hopefully indicates what I am trying to say.\footnote{The importance of giving the mathematical formalism a physical or empirical interpretation was stressed also in (Cao 2003) and (Lyre 2011). The defenders of OSR are often accused of trying to describe theories purely mathematically. Given the exposition in (Ladyman & Ross 2007) I do not think this is true, but they do come dangerously close to such a view.}

But string theory has unfortunately been quite removed from experimental input. The formalism has been ‘flying freely’ to a certain extent and this has caused some confusion. For instance it is very difficult to immediately distinguish the purely mathematical from the physical content. For this reason I will discuss the use of the word ‘duality’ in string theory.

6.3.1 On how the word ‘duality’ will be used

The word ‘duality’ is not new in physics or in mathematics. It has been used for quite some time and in many contexts. The focus in this article is on how the word ‘duality’ is used in the context of string theory. While there is some similarity to other contexts where the word ‘duality’ is used, such as in discussions on wave-particle duality in quantum mechanics, there are also important differences. To clarify, a few suggestions on how to restrict the use of the word ‘duality’ will be given.

A few things that must be distinguished from dualities in this context are the following:

1. It is well known that many equations and mathematical formalisms appear in different contexts describing completely different physical systems. For example the wave-equation and Poisson’s equation. These are not examples of dualities since it is clear that we describe different
empirically distinguishable systems. What the equations are used to describe are then obviously different. Hence, even if there is a mapping, isomorphism or structural similarity at the purely mathematical level it is not a duality since the application is not the same at the empirical level.

2. If it is clear that we are just discussing two alternative coordinate descriptions of one and the same system it will not be considered to be a duality. Hence different choices of gauge or coordinates should not be thought of as dualities. This is however something that might need further clarification. It might be the case that some dualities could be understood just in terms of a change of coordinates, but this should at least not be obvious from the start.

I believe that most physicists would agree that these are suitable restrictions for the use of the word ‘duality’ in string theory.

It is when we talk about variables and parameters that are not given an empirical or at least physical interpretation that we get into a more problematic situation. If we do not want to adhere to some form of strict operationalism/positivism we will think that we can meaningfully talk about alternative situations even on a more theoretical level. A physical theory is formulated using a combination of mathematics, physical interpretation of empirical parameters and analogies or metaphors from other linguistic practices to convey what we want to say. The question is which importance we attach to our way of describing the theories.

6.3.2 Purported dualities outside of string theory

Here I will briefly discuss two ‘dualities’ taken from (Zwiebach 2004). There they are introduced as introductory pedagogical tools to prepare the reader for the dualities in string theory. Since they are introduced mainly for pedagogical purposes, to prepare the readers for the dualities in string theory, I do not claim that Zwiebach consider these examples to be dualities of the kind that appears in string theory. The purpose of the discussion that follows is to clarify the criteria for relevant dualities in string theory that I gave above. I think that an explanation for why the examples are not dualities according to these criteria will be instructive for the reader.

**The harmonic oscillator**

The first example is a simple harmonic oscillator consisting of a mass $m$ hanging from a spring with spring constant $k$. The Hamiltonian is,

$$H(m,k) = \frac{p^2}{2m} + \frac{1}{2}kx^2.$$  \hspace{1cm} (6.1)

This harmonic oscillator has angular frequency $\omega = \sqrt{k/m}$ as can be deduced using elementary physics.
Zwiebach then suggests that there is a ‘duality’ transformation changing the parameters as,

\[ (m, k) \rightarrow \left( \frac{1}{k}, \frac{1}{m} \right). \]  

(6.2)

Neither the Lagrangian nor the Hamiltonian is preserved when this mapping is made, but the equations of motion for \( x \) is preserved i.e \( m\ddot{x} = -kx \). The parameter \( x \) indicates the position away from equilibrium. He shows that given a canonical transformation of the new Hamiltonian we can get back to the form of the old one. However, in a canonical transformation, we actually talk about different new variables with different interpretations so this does not change the fact that we have two situations that can be distinguished.

There is, of course, nothing wrong with the calculations in terms of the mathematical derivations. But if we assume a physical interpretation, and express the results in physical units, then the transformation only refers to the numerical values. For instance what is said is just that the oscillation of the coordinate \( x \) and the angular frequency \( \omega \) are the same in two different and empirically distinguishable situations. Hence, it should not be considered to be a duality according to the first criterion given above.

The two situations are for instance: when we have a mass of 1kg and a spring constant of \( 3N/m = 3kg/s^2 \) compared to when we have a mass of \( 1/3kg \) and a spring constant of \( 1kg/s^2 \). These situations are different and empirically inequivalent even though the equation for \( x \) would be the same.

If physics is supposed to be connected to empirically accessible results then we cannot talk about this as a significant example of a duality, at least according to the above formulated criteria. The reason for this is that not only the value of \( x \) is empirically accessible but also \( k \) and \( m \). It is important not to accept this example as an example of a duality unless we risk to make the term ‘duality’ too general as to make it more or less uninteresting.

As has been argued above, for there to be a relevant kind of duality we are not allowed to consider a duality to exist between two systems that are distinguishable at the empirical level, otherwise the point with dualities is lost. On the other hand the purpose with the example is pedagogical and is introduced to give students a simple example of similar calculations as the one they will encounter in the dualities in string theory. I do not claim that Zwiebach himself is not aware of this, he would probably agree with the point I have made. Nonetheless I do find it important to mention this example to illustrate an important point namely how we must not forget to distinguish between a purely mathematical formalism and its physical interpretation.
Classical electromagnetism

The second example taken from Zwiebach is Maxwell’s equations in the absence of sources:

\[
\nabla \cdot \vec{E} = 0, \quad \nabla \times \vec{B} = \frac{1}{c} \frac{\partial \vec{E}}{\partial t}
\]

\[
\nabla \cdot \vec{B} = 0, \quad \nabla \times \vec{E} = -\frac{1}{c} \frac{\partial \vec{B}}{\partial t}
\]

These equations are invariant under the following transformation,

\[
(\vec{E}, \vec{B}) \rightarrow (-\vec{B}, \vec{E})
\]

The point is made that this symmetry is not existent in the Lagrangian since we get a change of sign but still the equations of motion are the same.

But is this not just a possible change in terminology made possible from the symmetries in the equations without any substantial importance? If we have electric charges and no magnetic monopoles we can empirically decide which fields are to be interpreted as electric. That is we can decide which labels, \(E\) or \(B\) to assign to represent electric fields.

If there are magnetic monopoles the symmetry can be extended to include both magnetic monopoles and electric charges. The mapping is more a question of conventionality in terminology and empirical interpretation than anything substantial. It seems more like we have one and the same situation and the question is what we decide to call what. Hence, this should not be considered to be a duality due to the second criterion above. I would like to point out that by saying this I do not mean that this symmetry is uninteresting or that we should ban the use of calling this a ‘duality’. It is very interesting, but I do find the dualities in string theory to be more profound. They seem on the face of it to be more extreme and not immediately understandable in terms of a purely conventional choice made between interpretations of the mathematical symbols.

6.4 Dualities in string theory

In what follows a brief and rather nontechnical description will be given of string theory dualities. The point with these examples is not to explain the calculations or physical assumptions in any detail, the point is just to show what differs between dual descriptions.\(^\text{10}\)

\(^\text{10}\)For more details see (Polchinski 1998), (Zwiebach 2004), (Becker et al. 2007), and references therein. In the following I have also included references to a few important early research articles and review articles on the topic of dualities. For literature more directly addressed to philosophers see (Rickles 2008b) and (Witten 1996). The article by Witten has been republished in (Callender & Huggett 2001).
A number of different dualities have been formulated within the framework of string theory such as T-duality and S-duality. A more general version of T-duality is called mirror symmetry.\textsuperscript{11} Another kind of duality that has received a considerable amount of attention is the AdS/CFT correspondence.

These dual descriptions, if they are understood in a straightforward or literal way, present views of the world containing different kinds of objects and different descriptions of spacetime which can even differ in topology. Nevertheless they are thought to describe the same physics. This means that they give rise to the same set of particles, symmetries, scattering amplitudes and other empirically measurable, or at least potentially empirically measurable, quantities.

Some of these purported dualities are not rigorously proven mathematically. They are rather conjectures which most string theorists believe to be true. There are many mathematical tests of these conjectures and a host of different calculations have been done.

\subsection{6.4.1 T-duality and mirror symmetry}

When one of the extra dimensions used in string theory is compactified as a circle something surprising happens. It turns out that physically there is no real difference between a very small circle and a bigger one. If the radius in one solution is $R$ this can not be distinguished from a solution with radius $\frac{l_s^2}{R}$ where $l_s$ is the so called string length.\textsuperscript{12}

How is this equivalence supposed to be understood? Let me first try to explain how the mass-spectrum will be the same for the two different radii. Observe that the moving string is supposed to be quantized. Since the dimension is a circle we get quantization conditions deciding which quantum states are possible. This would be the same even if we had a point particle and not a string. Roughly the ‘wave-function’ of the string must obey periodic boundary conditions since the coordinate $X \equiv X + 2\pi R$. For large values of $R$ the allowed energies lie close to each other. Let these states be described by the quantum numbers $n$. When we have strings there will also be another effect. A string may be wound many times around the circular dimension. An analogy can be made with a rubberband wound many times around a rod. The larger $R$ is, the more energy is required to stretch the string many times around the compact dimension. Let $w$ be the winding-number that tells us how many times the string is wound around the circular dimension. The mass spectrum

\textsuperscript{11}This must not be confused with the more familiar kind of mirror symmetry, where things are symmetric with respect to the inversion made in an ordinary mirror.

\textsuperscript{12}An early article on T-duality is (Kikkawa & Yamasaki 1984), in this article the expression ‘T-duality’ is however not used.
for a certain radius $R$ is given by,

$$M^2(R; n, w) = \frac{n^2}{R^2} + \frac{w^2 R^2}{l_s^4} + \frac{2}{l_s^2} (N_\perp + \bar{N}_\perp - 2). \quad (6.3)$$

It is the two first terms that are of relevance here. $N_\perp$ and $\bar{N}_\perp$ depend on the internal vibrational state of the string and the constraint $N_\perp - \bar{N}_\perp = nw$ must hold. If the following exchange of radii is done,

$$R \leftrightarrow \frac{l_s^2}{R} \equiv \tilde{R}, \quad (6.4)$$

Then we get the spectrum,

$$M^2(\tilde{R}; n, w) = \frac{n^2 \tilde{R}^2}{l_s^4} + \frac{w^2 \tilde{R}^2}{l_s^4} + \frac{2}{l_s^2} (N_\perp + \bar{N}_\perp - 2) \quad (6.5)$$

So we see that,

$$M^2(R; n, w) = M^2(\tilde{R}; w, n). \quad (6.6)$$

This may convince us that the mass spectra are the same for the two dual radii. But it does not tell us that the theories are indistinguishable in all respects, since *prima facie* the radii of the compact dimensions are very different. However, there is a possible reinterpretation of the mathematical formulae such that a large dimension again appears instead of a small radius. I claim that a necessary precondition for this to be acceptable is that we do not have any *independent* empirical interpretation of the radii. This precondition is fulfilled here, in contrast with the example concerning the variables used to describe the harmonic oscillator, that was given above. 13

This is interpreted by saying that there is a minimal radius. This is due to the expectation that a classical description of spacetime is not reliable at the Planck-length scale. This does however show that even if we in the mathematical formalism first choose a radius well below the Planck length this radius is not to be interpreted as describing physical space.

When we are dealing with superstrings T-duality also changes the kind of string theory, so a type IIA string theory is mapped to a type IIB string theory and vice versa.

An even more radical form of duality is the so called mirror symmetries. In these the extra six dimension are compactified in such a way that they form examples of a specific kind of manifold called Calabi-Yau manifolds. It turns out that pairs of manifolds $\mathcal{M}_1$ and its mirror $\mathcal{M}_2$ that are *topologically* different nevertheless give rise to the same physics. When the underlying manifold is switched the type of string theory is also switched just as in the case with

---

13The above equations can be found in standard textbooks such as (Polchinski 1998) or (Zwiebach 2004), the choice of notation differ slightly between the different texts.
T-duality, so a type IIA string theory on $\mathcal{M}_1$ is equivalent to a type IIB string theory on $\mathcal{M}_2$.\(^{14}\)

From this it follows that according to string theory the description of the geometry and topology of spacetime can differ in formulations that are thought to be physically equivalent. However, it should be noted that in the case with the ordinary T-duality we found that only one of the dual radii that is allowed to be given a physical interpretation.

### 6.4.2 S-duality

In S-duality different formulations of string theory with different coupling constants are found to be dual. The coupling constant decides the strength of interaction between the strings and is not really a constant but depends on the specific solution. If $g_s$ is the value of the coupling constant there can be a dual theory with coupling constant $1/g_s$. The type IIB string theory is dual to itself under S-duality. The heterotic SO(32) string theory is S-dual to, the type I string theory.\(^{15}\)

One interesting result that appears here is that what seems to be fundamental objects in one formulation gets mapped to composite objects in the other. A composite object consisting of many strings, will in a dual formulation be treated as consisting of one fundamental string and vice versa. This means that what is treated as fundamental building blocks would be dependent on the description.

### 6.4.3 AdS/CFT

In the AdS/CFT correspondence a string theory is supposed to be dual to a quantum field theory defined on a different number of dimensions.\(^{16}\)

The expression ‘AdS’ stands for ‘anti-de Sitter’. An anti-de Sitter space is a space of constant negative curvature with Lorentzian signature. The expression ‘CFT’ stands for ‘conformal field theory’. A conformal field theory is an ordinary quantum field theory, with point particles and not strings, that is invariant under conformal transformation. The conformal field theory that is dual to the string theory that is defined on the AdS-space is defined on the ‘boundary’ of AdS.\(^{17}\)

---

\(^{14}\)An early paper on mirror symmetry is (Greene & Plesser 1990). A useful review article discussing both T-duality and mirror symmetry is (Giveon, et al. 1994).

\(^{15}\)The self duality of the IIB theory was established in (Hull & Townsend 1995). For the duality connecting heterotic SO(32) to type I string theory see (Polchinski & Witten 1996).

\(^{16}\)The seminal paper on AdS/CFT is (Maldacena 1998), which has inspired much further research in the area and a huge number of paper have been published. One important review article on AdS/CFT is (Aharony, et al. 2000).

\(^{17}\)This description is somewhat sloppy, the space in which the strings live is not just an AdS space, the whole manifold is the product of an AdS space and a sphere. Also the ‘boundary’
The AdS/CFT is an example of a holographic theory. In general this means that a theory in Y dimensions is equivalent to another theory in X dimensions. So if this is the case, the number of dimensions is dependent on the formulation. Also the kind of theory and whether or not we talk about strings or ordinary particles is formulation dependent.

6.4.4 Some general remarks on the dualities

Dual descriptions apparently describe very different physical situations. The geometry and topology including the dimension of spacetime can differ between dual descriptions. Which objects that are thought of as fundamental is also formulation dependent.

Then how are we to understand the dualities in string theory? We have theoretical formulations or descriptions of physical systems that seem very different but nevertheless result in the same physics. Here the word ‘physics’ is used by physicists in a way so that it refers to what can at least potentially be empirically testable. This is of course a somewhat questionable way of defining physics and it should be noted that there is a clear positivistic/instrumentalistic tendency in this way of speaking.

In (Zwiebach 2004, p. 386), we find the following:

Duality symmetries are some of the most interesting symmetries in physics. The term “duality” is generally used by physicists to refer to the relationship between two systems that have very different descriptions but identical physics.

It is interesting to note that he writes ‘two systems with different descriptions’ and not one system under different descriptions. Well what does he mean, are there two systems or one? If they were different systems no one would be surprised that the descriptions are different, but after all they have identical physics. So on what basis are we to believe that there really are two systems?

6.5 Different views on string theory dualities

Various interpretations of the dualities in string theory are possible:

- Interpretation 1.
  Accept the different dual descriptions as describing two different situations. If this view is taken we have a clear example of underdetermination. That is, the world may in reality be more like one dual description than the other but we have no empirical way of knowing this. This way

---

is not strictly speaking a real boundary but the conformal boundary, this is however of no real importance for the argument.

76
of looking at the situation entails that we are faced with some form of underdetermination.

Given this interpretation there are two alternatives:

– Interpretation 1A.

The two descriptions have the same empirical content, or at least potential empirical content, but besides that we can not say that they have an important X in common, where X could be a shared structure. If we accept this alternative we have an example of real underdetermination. This suggests that we can not know which, if any, theory or theory formulation it is that describes our world more accurately. This means that we must accept epistemic anti-realism since in this situation it is hard to find any reason for preferring one alternative before another.

– Interpretation 1B.

We accept them as two genuine alternatives that have an important X in common, where X could be a shared structure. This leads us to a position compatible with some weaker form of realism. If we can explicate the notion of structure and claim that there really is a shared structure this would be some form of ESR.

• Interpretation 2. We do not accept that they really describe different situations; instead they are descriptions of the same underlying reality which is given in terms of X. We might of course still accept the heuristic value of the alternative descriptions and the different ‘pictures’ used but we do not take descriptions concerning them literally. If this reality is purely structural we see the situation in the way a defender of OSR would.

The argument and possible interpretations given above can be seen as a special example of the general phenomenon where one can avoid epistemic anti-realism by dropping semantic realism or alternatively if one wants to keep semantic realism one needs to drop epistemic realism.

In the following it will be assumed that ‘X’ stands for ‘structure’. If we can find an alternative X we would consider other forms of intermediate realism than structural realism. Similar arguments to the following on structure could then mutatis mutandis be given.

Comparing views on structural realism; the differences between Cao, who defends a form of ESR, and the defenders of OSR can be described as follows. Cao would prefer Interpretation 1B while the defenders of OSR would prefer Interpretation 2. So the main difference is about the question about what is meaningful to take literally in a theory formulation. By this we see that there is a difference at the level of semantics and the proponents of OSR have taken a step in the positivist direction.18

---

18This is based on my reading of (Cao 1997, Cao 2003), (French & Ladyman 2003) and (Ladyman & Ross 2007). And their differing views on how to interpret quantum field theory.
Ladyman and Ross argue for OSR.\textsuperscript{19} Like the logical positivists they argue for a kind of verificationism, but their verificationism is not a theory about meaning, as the positivists version was. Instead their verificationism concerns what would be relevant to talk about in a scientific context.\textsuperscript{20} Since they describe their verificationism not in terms of meaning but what is scientifically relevant they do not think that their version of verificationism will face the same problems as the traditional logical positivistic version. Given this they would have to say that interpreting the different dualities as describing different scenarios is not just jibberish and empty words, but they would not endorse any such interpretation. For this to be consistent they need to accept different dual descriptions as meaningful but literally speaking false; that is if they want to distinguish themselves from an epistemic reading of structural realism. It is important to note that their verificationism, even though it is not a theory of meaning, still have consequences for how we should understand physical theories and that it leads to a non literal understanding of large chunks of the theory.

Physicists do usually not consider dual formulations as describing different genuine physical situations; so they would, if pressed, choose interpretation 2 just like the defenders of OSR. Note one important consequence of this: Since they claim that dual formulations are just descriptions of the same reality they can not then claim that literally everything in the formulations refer to something in physical reality. If it is the same underlying reality that is described one time with one manifold and another time with another topologically inequivalent manifold we can not take these statements at face value. This means that the large ‘old’ 4 dimensions epistemically and semantically receive a different status even though they are treated in the same way as the extra dimensions within the mathematical formalism of the theory.\textsuperscript{21} This is because they do receive an empirical interpretation, thus constraining what weird mappings we are allowed to do within the mathematical framework and still claim that we ‘describe the same thing’. What I would like to point out is that we have a complicated mathematical formalism which is only loosely connected to a physical interpretation. I would like to stress the importance of distinguishing between the mathematical formalism itself and the physical interpretation.

String theorists should say that the world is not literally like any of the dual descriptions in string theory as long as they defend a position similar to OSR and say that dualities describe the same underlying reality. For instance if one accepts formulations, with prima facie topologically different spacetimes in the mathematical formalism, as descriptions of the same reality; one must find

\textsuperscript{19}(Ladyman & Ross 2007).
\textsuperscript{20}In (Ladyman & Ross 2007, pp. 29-30), they describe their version of verificationism in further detail. They describe it as more resembling the view of Peirce than the view of the logical positivists.
\textsuperscript{21}This is discussed further in the following chapter.
claims that the physical spacetime has any of the suggested topologies to be literally speaking false.

It is important for any version of structural realism to carefully explain how their position differs from a view in which a theory is only determined by its empirical content. To do this we must further clarify what we mean by ‘structure’ in this context. It has been problematic to give a clear definition of structure. We cannot just use a simple definition of structure from logic and mathematics. I think one needs to carefully look at many different examples from actual physics and the dualities in string theory would be one such example. Any attempt to define structure in purely mathematical terms is bound to fail for reasons given above. The practice of making a physical interpretation of a theory and how it is supposed to connect to the empirical world must be made.

Note for instance the case that in T-duality the interpretation of the parameter $R$ as a very small radius below the Planck-length was disqualified, but this conclusion was not made solely on the basis of the mathematical formalism but depends heavily on a general understanding of physics. Note also that at the purely mathematical level different formulations might not even share the same structure there might be surplus structure in the formulation that is convenient but should not be thought of as representing anything physical, for examples and discussions of this see (Healey 2007) and (Lyre 2011). It might be that another kind of intermediate realism than structural realism would be the best way to understand our theories in physics.

6.6 Conclusions

What conclusions can be drawn from the discussion above? One thing is that it is not immediately clear how one should understand or interpret the existence of dualities in the context of string theory. The questions seem to be answerable only if certain semantic issues are clarified.

Physicists argue about string dualities in a way that seems most compatible with some form of structural realism (or at least some form of intermediate realism). Since most string theorists do not seem to think that dual descriptions give rise to real alternatives they are closer to OSR. They choose a specific formulation based on pragmatic reasons. They might for instance find certain ‘pictures’ to be heuristically valuable for certain purposes. The most important reason why a specific formulation is used is to allow for perturbative calculations. Depending on the state of the system one formulation might be tractable while another might not. This suggests that in certain situations one description is better than another. Sure this is the case in terms of how easy we can perform calculations. But can it be said that for a certain state the tractable formulation gives a better description of the underlying reality? Of course not, and what are we to say when the state is such that both formulations can be
used? I think it is fair to say that the common view among physicists concerning the dualities is such that it cannot be understood as endorsing a view where everything in a theory formulation is taken to be literally true, not even tentatively.

While most physicists would understand the dualities in a way similar to OSR, alternative views are possible. The different views depend on the semantic question regarding what to take literally in a theory formulation.

I think that semantic issues concerning theories are important and should be further discussed. After the fall of positivism that endorsed a very strict empiricist semantics I think that many philosophers of science have treated this question too lightly and just assumed that any reasonably understandable expressions has a clear meaning. It is important to understand that when scientists formulate and develop new theories the meaning of what they say is not immediately clear. It is fair to say that the physicists sometimes do not really know what they are talking about and a physical interpretation cannot be read directly from the formalism. This is however not meant as criticism; physicists who develop new theories understand that their present formulations at best only give us an approximate description and also that even the reference of their terms can be quite unclear. Given this I think it is important to further readdress these questions and discuss the difference between a mathematical formalism and its physical interpretation. I believe we adopt some form of empiricist semantics, which avoids the pitfalls of logical positivism. The way I have discussed the different ways in which we can view dualities in string theory is a step in this direction.

When describing theories and what they mean and how they are used there is one point that I find very important to make. That is, that even if one denies any deeper importance to the 'pictures' and 'images' that a specific formulation of a theory conveys, I think it is very important to not try to avoid the use of such pictures. There is an undeniable heuristic value of an interpretation containing pictures. Different pictures might suggest different possible extensions of the theory or inspire new hypotheses. So regardless of which view to take on the semantic issues it is imperative that this should not be used to argue that we must formulate a theory in a pure form which only contains what is deemed completely acceptable according to the standards of the semantic theory. We do not want to put the physicists in a semantic straitjacket.

The question on realism and anti-realism is important for physicists to consider. There is a tension between realist and anti-realist tendencies. When it comes to the dualities in string theory, the standard view among string theorists implies that they deny semantic realism, and I think rightly so, but in other situations string theorists have opted for a very realist understanding. For instance in the debate on the so called ‘Landscape’ of string theory, where different solutions to string theory are considered, some string theorists think that we
should interpret this as different parallel and really existing universes.\(^{22}\) I will not discuss the rather controversial debate concerning the ‘Landscape’ in this chapter any further, I only want to point out that string theorists need to perform a rather delicate balancing act between realism and anti-realism in their attitude towards their theoretical constructions. Some form of intermediate realism such as a version of structural realism seems to be recommended. The physicists view seems to be most compatible OSR. I also think that something akin to OSR is the best way of understanding the dualities. However, as has been argued above I harbour some doubts as to whether ‘structure’ can be explicated in a satisfying manner or whether or not ‘structure’ is the best word to use to describe a defensible kind of intermediate realism concerning physical theories.

That the dualities of string theory strongly suggest something akin to structural realism was also argued in (Dawid 2007). He also explains that the dualities can be used to undermine a view that takes the specific ontologies of the dual descriptions seriously. In this chapter I have described in more detail how the arguments behind these conclusions are connected with certain semantic issues concerning the physical interpretation of theory formulations that rely heavily on mathematics and are not directly tied to observable data.

I will finish this chapter with another lesson that I think that we can learn from the dualities in string theory. If it is possible to find dualities between seemingly very different formulations within one research programme one should acknowledge that it might not be suitable to only study one research programme. The hostility that from time to time has appeared between proponents of different research programmes in quantum gravity might be a serious mistake. If something like a duality would be found between the research programmes, then they could actually merge.

This provides us with a new reason for embracing the general methodological suggestion given already in chapter 4. Namely, that when no theoretical research programme produce empirical results then a broader perspective ought to be adopted. I think that only if a research programme is empirically progressive, would it be rational to completely ‘stick with the programme’.

\(^{22}\) For a popular description of this view see (Susskind 2005).
7. Quantum gravity semantics and the nature of spacetime

This chapter is based on the manuscript (Johansson & Matsubara submitted). It has however been modified so as, to a slightly larger extent, also address other research programmes in quantum gravity than string theory. It can be seen as a continuation of the argument put forth in the previous chapter. To avoid unnecessary repetitions I have omitted the discussion on dualities which is treated in much more detail in the previous chapter.

7.1 Introduction

Theories of quantum gravity, like other highly abstract physical theories naturally raise questions of interpretation. What is the meaning of theoretical concepts like string, brane, manifold, field, loop etc.? Should we take all concepts being used in those theories as having physical significance, or are some mere calculational devices, devoid of physical meaning? What does a theory such as string theory, tell us about reality, if it is true? That is the central question in this paper. While the focus is on string theory, I think that the conclusions in this chapter are more general and apply to physical theories in general when they become highly abstract and far removed from our everyday experiences of reality.

I will continue to use the example of string theory to argue for an empiricist position in the semantics of modern theories of physics. My position is, in outline, that one should make a clear distinction between physical and mathematical claims in theoretical physics and thus, by implication, I indicate that this distinction is not sufficiently upheld in the current discussion. Secondly, it will be argued that the concepts used in saying things about the physical world must have a connection with concepts used in observation reports. This view will be fleshed out by discussing some aspects of string theory.

The here presented position deviates both from classical positivism, according to which theoretical statements lack truth values. It also deviates from van Fraassen’s constructive empiricism, according to which they have truth values but are grossly underdetermined by any possible amount of observations.

In this chapter I will in particular critically investigate the concept of dimension of spacetime, as it is used in string theory. It is commonly claimed that if string theory is true, or approximately so, there are 10 dimensions of
spacetime, of which six are compactified, three are the usual dimensions of space and one is time. The question is: What does it mean to say that there are in fact more dimensions of physical space than we observe?

The core issue is thus the relation between the mathematical formalism of string theory and its physical interpretation. I will then present a more general discussion concerning the problems of giving physical interpretations of mathematical formalisms.

To avoid any misunderstanding, I point out that it is a trivial point that in mathematics one can have spaces with as many dimensions as one would like. Common examples of the kind of mathematical spaces which can be encountered in physics are Hilbert spaces, Riemannian and symplectic manifolds etc. The phase-space of $N$ particles in 3-dimensional space is for instance $6N$-dimensional. I take all of this for granted and assume that the number of dimensions in this mathematical sense is understood and unproblematic. As the example with the particles clearly illustrate even if a mathematical space is given a physical interpretation it does not mean that these are the dimensions of physical space or spacetime. I stress this, even though it is very elementary, to avoid some misunderstandings which I have encountered when previous versions of this text has been commented upon. What I am discussing in this text is the number of dimensions specifically of physical spacetime.

In the previous chapter it was argued, based on the dualities of string theory, that we can not take all statements in a theory formulation literally. The consequences of the dualities in string theory for the claims concerning the number of spatial dimensions was not addressed in much detail there. In this chapter I take Interpretation 2. in the previous chapter for granted.

The conclusions drawn in the previous chapter, by looking at the dualities of string theory, were that neither the specific metric, nor the topology of the extra dimensions, including the number of dimensions, can be given a straightforward physical interpretation. Of course, in practice, one needs to use one formulation or another. But the existence of these dualities show that the use of a particular manifold that *prima facie* suggests that spacetime has a certain number of dimensions, topology and metric can not be taken at face value.

7.1.1 From classical to quantum

String theory is a quantum theory and, as is typically the case with such theories, it is first formulated as a classical theory which is then quantized. The classical theory involves a mapping of the worldsheet of strings into a target space manifold. The target space is what one often thinks of as representing as physical spacetime.
When the world sheet is quantized we get a conformal field theory in two dimensions.\(^1\) In a sense, the two-dimensional conformal field theory on the world sheet, is more fundamental than the target manifold in the theory. This is to some extent a key to understand dualities. Witten writes the following concerning mirror symmetry:

\[
\text{...which is a relationship between two spacetimes that would be quite distinct in ordinary physics but turn out to be equivalent in string theory. The equivalence is possible because in string theory one does not really have a classical spacetime, but only the corresponding two-dimensional field theory; two apparently different spacetimes } X \text{ and } Y \text{ can correspond to equivalent two-dimensional field theories.}
\]


Note that he has written “in string theory one does not really have a classical spacetime, but only the corresponding two-dimensional field theory”. This understanding of string theory is today generally accepted, but what it entails concerning the physical status of the extra dimensions has not been explicitly discussed as much as I think it should be.

One should here observe the plural of the word ’spacetime’. Is Witten here talking about a number of different physical worlds, with different spacetimes, or is he talking about a number of different possible mathematical representations of the same physical spacetime, the universe? I think that it is the second option that is by far more reasonable. In other words, ’spacetimes’ must be viewed as denoting mathematical objects, not physical ones, and likewise the indefinite expression ’a spacetime’ means a mathematical object. But sometimes the word ’spacetime’ is used as name for the spatiotemporal structure of our universe, whatever that is. In what follows, I will use the expression ’physical spacetime’ for the latter.

The fields on the two dimensional world sheets in string theory are normally interpreted as specifying the position of the string in the target manifold, which is supposed to represent physical spacetime. However in the general formalism of two dimensional conformal field theory this interpretation is not at all mandatory. When two dimensional conformal field theory is used in other contexts such as statistical mechanics the fields are of course not thought of as representing physical spacetime.

Another thing, to keep in mind, when discussing the interpretation of the mathematical formalism of string theory, in terms of dimensions of physical spacetime, is the peculiar way in which heterotic string theory is formulated. Vibrations on a closed string propagating in one direction along the string are

---

\(^1\)Two dimensional conformal field theories have many special properties that are not existent in other number of dimensions. These theories have also been studied outside of the context of string theory and have relevance also within statistical mechanics. A very comprehensive account can be found in (Di Francesco, et al. 1997).
independent of those in the other. This makes it possible to treat one direction of vibrations as purely bosonic, thus preferring a 26-dimensional space to live in. The other direction of vibrations is treated in the supersymmetric fashion. For this to result in a theory in 10 dimensions, 16 of the dimensions for the bosonic direction must somehow be eliminated. This turns out to be possible in two different ways, resulting in the versions of string theory called Heterotic $SO(32)$ and Heterotic $E_8 \times E_8$ respectively. The extra 16 dimensions are then not interpreted as representing physical spacetime; they are taken to represent internal degrees of freedom.

It is commonly assumed that any quantum theory of gravity will radically change our understanding of physical spacetime. It is assumed that physical spacetime no longer can be treated as classical, i.e. as faithfully being represented by a continuous manifold, below the Planck-length. Hence the classical continuous description of physical spacetime, in terms of pseudo-Riemannian manifolds, is not to be trusted at these scales. So thinking about, quantum strings moving in continuous physical spacetime, is presumably not a correct picture of the real world at the fundamental level, even if one accepts that string theory is the best path towards quantum gravity. As long as completely different metrical and topological properties can be assigned to different versions of a formalism, all held to describe the same reality, we should not conclude that physical spacetime really has any of the variable properties.

If there is such a deeper more fundamental description, in which there is no continuous spacetime, (i.e. a continuous manifold with a metric) it is reasonable to reserve the word ‘physical spacetime’ to the part of the specific solution which in a sufficiently unambiguous, albeit of course approximate, way could be represented by a continuous differentiable manifold with a metric. This means that no conclusion about the number of dimensions of physical spacetime can be drawn from the features of the mathematical formalism only, since number of dimensions of physical spacetime depends on the details of the solution that is supposed to represent our world. Since such a solution is not yet found, such a conclusion would be premature. Note, that this conclusion does not rule out that there may be extra dimensions of physical spacetime.

### 7.1.2 Background-independence

As stated above the idea that a traditional classical spacetime must be abandoned at a fundamental level is common among many approaches to quantum gravity. A leading proponent of the rivaling loop quantum gravity approach has written:
The key conceptual difficulty of quantum gravity is... to accept the idea that we can do physics in the absence of the familiar stage of space and time. We need to free ourselves from the prejudices associated with the habit of thinking of the world as “inhabiting space” and “evolving in time”.
(Rovelli 2004, p. 10)

The reason why dualities appear in the first place may be the postulation of a background spacetime for the strings, a background which is described in classical terms as a differentiable manifold. Such a starting point is at variance with the fundamental aspect of the general theory of relativity, or GTR, viz, that physical spacetime is not something inert, but is a part of the dynamics. The interpretation of GTR is not beyond dispute, but no one disputes that spacetime is a dynamical entity.

Accepting that physical spacetime is a dynamical entity, forces us to the conclusion that a postulation of a non-dynamical spacetime background for string theory cannot be correct. This, of course, all string theorists accept, but they think that it is still a useful approximation and that further development of the theory will improve the situation.

But some physicists, with a philosophical bent, who support loop quantum gravity has claimed that a truly fundamental theory should fulfill the demand for background independence from the start.2 The notion of background independence is however ambiguous, see (Belot 2011). Belot’s conclusion is that whether a theory is background independent or not is not a mere formal feature, but depends on our interpretation. His explication is:

[a] theory is fully background independent if by its lights two solutions instantiate the same geometry iff they represent the same physical possibility and it falls short of background independence to the extent it features solutions with same geometry that represent distinct possibilities.
(Belot 2011, p. 2865)

String theorists do hope to be able to formulate a truly background independent formulation of string theory but this has not been done. As string theory now is understood, theory construction begins by formulating a classical, i.e., non-quantized theory, which then is quantized.3 And here a classical background is originally postulated and the calculations are perturbative around that classical background. The dualities are thought to give us hints on the non perturbative aspects of the theory and presumably throw light on the background independent formulation of the theory that proponents of string theory assume exists.

2See (Smolin 2006a), (Rickles & French 2006), (Rickles 2008b) and (Rovelli 2004, Rovelli 2007) for discussions on the importance of background independence in quantum gravity.
3In a weak sense string theory has some background independent features since it can be shown that string theory demands that the metric of spacetime satisfies Einstein’s field equations (with further higher order corrections).
Given the arguments, concerning the dualities, it is thus reasonable to assume that in a more fundamental description one should be able to start with a quantum theory and then directly formulate solutions. In this formulation we would not get dual descriptions describing different classical spacetimes.

When pushed, string theorists usually acknowledge that it is a problem that string theory is formulated as a background dependent theory, but they hope that in the end string theory will receive a background independent formulation, maybe in terms of a more developed understanding of M-theory. The view, held by proponents of string theory, is that there must exist a more fundamental background independent formulation of string theory, and I certainly do not rule this out. However such a formulation seems to be urgently needed to make sense of the dualities which have been discussed in this chapter.

It seems to be quite a general conclusion, shared by many approaches to quantum gravity, that the continuum of spacetime must be abandoned in a fundamental description of nature. Given that it is not that surprising that some notions based on an approximate scheme can not be taken at face value.

7.2 Physical content and physical concepts

Physical theories rely heavily on the use of abstract mathematics. But physical theories are intended to be about the physical world, not about mathematical entities. This raises semantic questions regarding the connection between the mathematical formalism on the one hand and the theory’s claims concerning physical reality on the other.

In the old days of logical positivism a strong empiricist view on semantics was espoused. The cognitively significant content of a theory was thought to be captured in its observable predictions. Sentences or propositions which could not be given an empirical interpretation were considered neither true nor false, they were viewed as mere auxiliaries for making predictions, hence the label ‘instrumentalism’.5

Today the logical positivists are thought of as being anti-realists concerning science, even though not all of them would be happy to describe themselves as such. The reason for this is that they restrict the scientific theories to the observable empirical content and refuse to speculate on metaphysical issues on what the world really is like behind the observable predictions. They found such questions meaningless and irrelevant for science, thus their semantics was closely connected with their epistemology. Very roughly; if we can not

---

4 Note for instance, as was mentioned in chapter 3., that both LQG and causal-set theory claims that spacetime actually is discrete at a more fundamental level.

5 This is of course a very simplified description of logical positivism which underwent a number of changes and developments during the years. For an excellent discussion on the evolution of logical positivism see (Suppe 1977).
explain what observations a theory entails we also do not know what we are talking about.

Nowadays questions concerning semantics and epistemology are typically treated as separate. Among modern scientific realists and anti-realists we can find agreement that the semantic question is to be treated in separation from the epistemic question. A modern proponent of scientific realism would take for granted that statements in science can be seen as meaningful even without or before we can test their empirical consequences in any way. Realists assume that statements being part of scientific theories have a specific truth value and that we ought to take the claims of theories literally. The most influential modern version of anti-realism regarding science, the constructive empiricism of van Fraassen, is in agreement with this view. They are hence both semantic realists. Their disagreement is whether or not we should trust the theories beyond what they say about observable states of affairs; van Fraassen claims that the only thing that matters is whether the theory is empirically adequate or not; its truth is not the issue, since both true and false theories can be empirically adequate. Scientific realists disagree; they claim that we have good reasons to infer the truth of successful modern theories, since their truth is the best explanation for their success.

I support some form of intermediate position in the debate on realism versus anti-realism concerning scientific theories. I hold that epistemological and semantic questions cannot be treated separately; in this I agree with the view of classical positivism. But I don’t think that verification, in the sense used by positivists where the applicability of individual terms are supposed to be given verification criteria. I further disagree with the strong interpretation of semantic realism, according to which all statements in a theory are either true or false, taken as statements about the physical world. The point is that in a highly abstract theory such as string theory, some parts may be viewed as mathematical structures that do not represent any physical structures. So these parts lacks physical significance. But they don’t lack mathematical significance. If interpreted as only displaying relations between mathematical concepts they are true (or false, if we have made mistakes). It is unfortunately often very difficult to determine what is physical and what is pure mathematics in a physical theory, especially when the connection to empirical results is weak.

The core issue is hence aboutness: which concepts in the theory are used for classifying physical phenomena and which are mere mathematical concepts? Any answer to this question is at the same time part of epistemology and of semantics.

So what is needed is a way of distinguishing between those parts of a physical theory that can be interpreted as being about physical reality and those that are mere mathematical auxiliaries. Fortunately, there are some well-known cases where this distinction has been drawn to most people’s satisfaction; so
without relying on any general principle, I have some paradigm cases that will guide us.\(^6\)

One such paradigm case is the difference between the phase of the wavefunction in quantum mechanics and its squared amplitude. The phase does not represent a real property of quantum systems,\(^7\) whereas the squared amplitude of the wavefunction does. Similarly, in quantum field theories there are gauge-symmetries; different choices of gauge do not correspond to different physical situations. Hence, there is a mathematical redundancy in our description of physical reality. It is generally accepted that if some aspects of a physical theory depend on the chosen gauge, they have no physical realisation; only gauge-invariant, or otherwise transformation-invariant features are candidates for being real properties.\(^8\)

Another well-known example is provided by the so-called ‘hole-problem’ in GTR. The upshot of this is that points, on a bare manifold, can not be thought of as representing points in physical spacetime independently of the metric defined on them. Mathematical manifolds which can be transformed into each other, via an active diffeomorphism, represent the same physical spacetime. This may be obvious to modern physicists, but once it was a hard-won insight and maybe still is to those not acquainted with GTR.\(^9\)

Using this kind of reasoning among physicists as a template, I can reason as follows. Since two dual versions of string theory represent one and the same underlying reality, as has been assumed in this chapter, I have good reasons to hold that if these versions have different topologies and/or different dimensions, these features have no physical meaning. It would be just as weird asking which one of the topologies in a dual description is the real one, as asking which gauge represents the real world.

So when a string theorist talks about a spacetime with a particular number of dimensions and a specific topology for which there is a dual description with another number of dimensions or topology I do not understand such talk as a literal description of reality, or of alternative realities that are possible according to the theory. I interpret the word ‘spacetime’ as denoting a mathematical structure devoid of physical interpretation. In my view, string theorists have not reflected sufficiently on these issues, often they do not seem to keep apart the mathematical and the physical meaning of ‘spacetime’.

---

\(^6\)In the cases here discussed the alternative is to accept that the extra mathematical structure do represent something physical. This view has been suggested by some philosophers. If this route is taken, underdetermination of one sort or another follows. For a discussion on various interpretative moves of this sort, see (Rickles 2008c). He himself does however support a minimalist structuralist understanding of the theories due to the existence of the many metaphysical options available. The description given here portray the standard view among physicists.

\(^7\)However phase differences are real properties.

\(^8\)For a detailed discussion on gauge theories from a philosophical perspective see (Healey 2007).

\(^9\)Einstein has commented upon his troubles with developing GTR: “The main reason lies in the fact that it is not so easy to free oneself from the idea that co-ordinates must have an immediate metrical meaning.” (Einstein 1951, p. 67).
This result is not that surprising if seen in the light of an empiricist semantics. To be fair, by taking for granted that the dualities describe the same reality I have to a large degree already assumed a kind of empiricist semantics. In answer to that, I can say that I just draw out consequences of the standard physicist interpretation. An interpretation which I in previous chapters have argued to be the most reasonable. My reason to reject the alternative view where dualities are thought of as representing genuine physical alternatives is due to my general view in favor of some empiricism in my view of semantics. This I started to argue for already in chapter 5. There the arguments were not based on the dualities in string theory.

To give another argument, in favor of empiricism in semantics, I will take as a starting point our concepts of body, space and time, as they are used in classical physics and in ordinary language. The point with this is not to claim that deviations from this way of conceiving space are forbidden or that the classical notions of space and time are sacrosanct. In fact I rather believe that they are deficient as a description of reality at the most fundamental level. The point is rather to clarify the role of classical notions in giving physical content to theories. For I see no alternative to the view that a description of any truly observable phenomenon must ultimately be given in terms of observations of events occurring at a certain place at a certain time. And observable events are basically bodies sending signals which we may observe. Hence, observation reports are ultimately given using the concepts of body (or some derivative one, like 'meter', 'pointer' ‘counter’, etc.), space (or spatial distance) and time (or time elapse). This is reflected in the SI system, in which the units for quantities attributable to bodies, time and space intervals, i.e., kg, second, and meter, are three out of the seven fundamental units.

In order to give empirical content to a physical theory, we must in some way connect concepts in such a theory to this classical realm, for observable phenomena must be described using classical concepts. This point was repeatedly made by Bohr when discussing the interpretation of quantum theory, and the point is generally valid. The argument is simple: From an empiricist point of view, a necessary condition for us to be able to say about a theory that it has empirical content is that it must have observable consequences. An observable consequence is something we humans can locate in space and time. Hence, no matter how abstract our theory is, classical concepts of spatial location and time, must be used, when describing its empirical consequences.

Observe however that this does not mean that I require that every concept used in theories must be defined in observational terms, for that would be a version of classical positivism. Nor do I assume that the classical formulation needs to be used in a more fundamental description of a theory; I only want

---

10One of Bohr’s formulations is the following: "...it is decisive to recognize that, however far the phenomena transcend the scope of classical physical explanation, the account of evidence must be expressed in classical terms." (Bohr 1951, p. 209).
to stress that in order to give physical content to a theory, some of its consequences must be put in connection with this classical realm. But this means that some theoretical concepts must be defined, at least implicitly and perhaps via a chain of intermediate concepts, in terms of concepts used for observation reports.\textsuperscript{11}

String theory is now unfortunately far removed from direct empirical observations. I would however not say that string theory is completely devoid of physical content, in the sense of making claims about the physical world. These claims follow from the way in which string theory is supposed to connect to other theories such as general relativity and quantum field theory. It is known, to a certain extent, what in string theory is supposed to correspond to what in the earlier theories. For instance what, in string theory, that is supposed to corresponds to Feynmandiagrams in QFT etc. This means that since this is an integral part of string theory, physical content is in a way inherited. However I think that the further removed from the classical concepts the less clear is it what a theory actually claims. Basically we understand less and less what we are talking about.

The development of a theory involves developing new concepts. When doing this scientists tend to use old words in new contexts where they can be quite misleading, they use analogies that are not meaningless but to different degrees vague and ambiguous. It is also the case that I do not think that we can define concepts in theories of these kind independently from each other; they depend on their place in the theoretical structure.

This view is also quite compatible with the way physicists reason. It is not uncommon to see statements to the effect that we do not yet understand string theory. This is also said about quantum mechanics. The idea with this kind of empiricist semantics is not to draw a sharp boundary between, well defined and clear empirical statements with well defined truth values, and purely theoretical statements without truth value in the style of the logical positivists. Instead the idea is to point out that the distinction between physical statements and merely mathematical statements becomes less and less clear the further removed we are from our everyday classical observations.

It should also be pointed out that I do not take for granted that two empirically equivalent theories always should be identified as being the same theory. Rather this depends on the specific case, whether to judge this to be the case or not. This can of course be difficult. But I think it is justified to identify dualities and different choices of gauge as representing the same reality. Different interpretations, of the foundations of quantum mechanics, can still be seen as genuine alternatives that differs by making different substantial claims about

\textsuperscript{11}A classical uncontroversial example is electromagnetism. The electromagnetic core concepts are electric field, magnetic field, charge and current. These are connected to the mechanical concepts force, mass, velocity and acceleration, by Lorentz’ law and Maxwell’s equations. These laws jointly make up implicit definitions of the fundamental electromagnetic concepts.

91
the nature of reality. While I think this is the case, the distinction needs to be further clarified.

7.3 Summary and conclusions

I have argued that, based on the present status of string theory, there is insufficient justification for saying that it demands a specific number of extra dimensions of the physical world. It is not even clear if talk about spacetime dimensions on a more fundamental level is justified. The different versions of string theory, assumed to all give the same physics, differ in regard to both topological and metrical properties; this strongly supports the view that such topological and metrical properties of the mathematical formalism do not represent aspects of reality.

I do not rule out the existence of extra dimensions of physical spacetime; I only point out that at present the physical status of the extra dimensions in string theory is uncertain. A more developed understanding of string theory, perhaps in the guise of a background independent formulation, combined with better empirical evidence, may be needed for saying how many dimensions our world really has.

A suggestion given above is that we should reserve the expression 'physical spacetime' for that part of a string theory solution that can be approximated in a sufficiently unambiguous way by a pseudo-Riemannian manifold.

I do not want to attack the use of the specific 'pictures' used in different dual descriptions. They are heuristically and pragmatically useful. But it is important to carefully keep in mind which parts of such pictures that have physical significance and which lack it, just as when we use different kinds of models for atoms and molecules.

Also it should be pointed out, that the general empiricist attitude argued for here, on the basis of string theory, also could be applicable to other approaches to quantum gravity. On the one hand it is quite plausible that something like dualities may be found within, or between, other approaches to quantum gravity. Also some of the arguments, in this and previous chapters, do not rely on the dualities of string theory.
8. The many worlds of physics and philosophy

This chapter is based on a manuscript, which at the time of writing this, is still work in progress.\(^1\) Due to this, this chapter is slightly more rough compared to previous chapters. There are quite a few arguments and other things that I would like to develop further before sending the manuscript to a journal. Having said that I certainly think that this chapter, even in this slightly rough state, deserves to be included in this dissertation.

### 8.1 Introduction

For various reasons, the thought that there may be more worlds than the one we inhabit, has appeared in a number of contexts throughout the history of philosophy and physics.\(^2\) The motivations for proposing multiple worlds differ widely between the different contexts in which they have been discussed. Here follows a quick presentation of the different discussions on multiple worlds that will be investigated in more details in the rest of the article.\(^3\)

Within philosophy the discussion on possible worlds has a long history, famous are Leibniz’s reflections on possible worlds where he argues that God chose the best of all possible worlds to be actualized. While this idea has been ridiculed by many, it should be noted that according to Leibniz the other possible worlds do not exist in the same way as our own. This is also the most common attitude towards possible worlds among philosophers. Possible worlds are thought to be possible counterfactual ways in which the world could have been, they are typically not thought to really exist in the same way as our own world. Among philosophers, Lewis (1986) has however famously argued that we ought to treat the possible worlds as equally real and existing in the same way as our own world. In this chapter, it is primarily this Lewisian approach to possible worlds, that will be discussed and compared with the multiverses suggested by physics.

In physics multiple worlds have been discussed for various reasons. Well known are the discussions on various many worlds interpretations of quantum mechanics.\(^4\)

---

1\(^{\text{(Matsubara work in progress, a).}}\)
2\(^{\text{Recently quite a lot has been written on different ideas on multiple worlds. Popular science books, such as (Susskind 2005) and (Greene 2011) and also more scholarly publications for instance (Carr 2007).}}\)
3\(^{\text{There are some other discussions on multiple worlds that I leave aside, such as discussions on computer simulated worlds.}}\)
4\(^{\text{The original idea behind the many worlds interpretation goes back to (Everett 1951).}}\)
More recently, considerations based on string theory, have suggested a new way in which there might be parallel worlds. String theory, as far as it is understood today, seems to allow for many possible solutions each of which would result in different worlds with different ‘laws of nature’. The scare quotes are intended. I want to indicate that the expression ‘laws of nature’ needs to be clarified. The ‘laws of nature’ here mentioned are not the fundamental laws, but rather effective laws applicable in a specific solution.

Originally the hope with string theory was that it would be able to predict a specific vacuum solution; thus explaining the values of the various parameters in the standard model of particle physics. This hope is now more or less abandoned due to the discovery that string theory seem to allow for an extreme number of different solutions. Instead of looking at this as a complete failure of string theory, a different attitude has been adopted by quite a few researchers. They argue that a huge number of these different solutions are physically realized. This way of arguing is closely connected to the quite controversial anthropic principle. The thought is that with a multitude of physically existing worlds the apparent fine-tuning of the parameters of our universe can be explained without the need of a creating deity.\footnote{See (Susskind 2005, Susskind 2007). A good review is (Schellekens 2008), especially the extended Internet version. There one can also find references to further original papers.}

Many, however are quite critical of the use of anthropic reasoning, finding it empty and unscientific. Critique against the use of anthropic arguments has for instance been given by (Smolin 2006b)(Smolin 2007). Noteworthy is that even though Smolin judges the anthropic argument to be unscientific, it is not the idea that there are multiple worlds or universes that he sees to be the problem. In fact he proposes another scenario which also involves a multiverse.

There are also further ideas involving the use of multiple worlds that will be discussed in this text. Such as Tegmark’s taxonomy of ways in which we can have multiple worlds.\footnote{See (Tegmark 2003, Tegmark 2007, Tegmark 2008).} The purpose of this chapter is to critically examine various claims that are involved in the different views that in various ways take the idea that there are other worlds, in one sense or another, seriously.

8.2 A few comments on terminology

First a few comments on terminology. I prefer to use the term ‘world’ instead of ‘universe’ but the terms will be used interchangeably especially since I need to quote different authors with different preferences. This is not of great importance, it is merely a terminological choice. One reason to avoid talking about multiple universes is that the term ‘universe’ traditionally is understood to include everything that physically exists. This is of course not a strong reason, we can start to use a term in a new way without problem as long as we are clear about it. That ’atom’ is derived from a Greek word meaning...
indivisible, does not force us to use it so that it refers only to indivisibles as we are well aware of, the situation here is analogous.

Another problem that arises is, that when Lewis talks about his independently really existing possible worlds, they are thought to be completely physically and spatiotemporally disjoint. On the other hand in physical scenarios, whether we discuss the Landscape of string theory, the many worlds of quantum mechanics or far removed patches in one universe they typically are physically connected in one way or another. While ‘worlds’, in this sense, are typically largely autonomous and do not interact to any large extent they are not as clearly separate and non overlapping as assumed in Lewis account of possible worlds. This suggests that it is instead more suitable to say that we are dealing with different, fairly independent, parts of one world.

Despite this, I think that some ideas from the philosophical considerations on possible worlds can fruitfully be used to elucidate more physically motivated thoughts and speculations concerning what they call ‘worlds’. Here the word ‘world’ will thus be used in a rather loose sense, so as to be able to embrace all the different discussions in philosophy and physics. More precise distinctions will be introduced when needed.

8.3 Possible worlds in philosophy

Possible worlds have been a topic of interest in philosophy for a long time. Nowadays talk about possible worlds in philosophy is intimately connected with modal logic.

Modal logic deals with the concepts of possibility and necessity. The idea of discussing possibility and necessity using modal logic, received better foundations with the work of Kripke using possible worlds semantics. Using Kripke semantics the consistency of different systems of modal logic can be shown. This is however a purely formal or mathematical feat. The ‘worlds’ used are defined in terms of abstract set theory and no further interpretation needs to be given to these things at all. It is fair to say that it only provides us with a logical skeleton. Questions concerning the physical or metaphysical status of these worlds are not answered by the formal apparatus of modal logic.

The use of modal logic is common and widely used, perhaps overused, among analytic philosophers. Modal logic is, however, not well known among physicists. Since I want this text to be accessible to both physicists and philosophers I below, very briefly, explain the basics of this formalism.

One starts by assuming that there is a set of possible worlds which is denoted \( W \). An accessibility relation \( R \) is postulated. Given this relation \( R \) and \( w_1, w_2 \in W \), the expression \( w_1 R w_2 \) states that \( w_2 \) is accessible from \( w_1 \). Then if \( p \) is a proposition we can state the proposition \( \Box p \) which is typically supposed to mean that \( p \) is necessary. This proposition is true in world \( w \in W \), if

---

7 See (Kripke 1959).
and only if, it is true in all other worlds in $W$ that are accessible from $w$. In the same manner the proposition $\Diamond p$, which is typically supposed to mean that $p$ is possible, is true in $w$ if it is true in some world accessible from $w$.

The properties of the accessibility relation decides which kinds of modal logic that will be used. For the purpose of this paper I do not need to go into further details concerning this.\footnote{For details see any standard textbook on modal logic, such as (Hughes & Cresswell 1968).}

If the accessibility relation is such that every world in $W$ is accessible from every other world in $W$, we get the traditional description of necessity and possibility. That is, necessity is truth in all possible worlds and possibility is truth in at least one possible world.

But how are we to interpret this formalism? As a formal calculus it is quite unproblematic, but seen purely abstractly there is no reason to see the formalism as describing real worlds, or that the box and the diamond describe necessity and possibility in any relevant sense. Other interpretations of the symbols can be made.

To give the formalism an interpretation we need to decide what the set of worlds is and the content of the accessibility relation. We need to have some specification of our use of possibility such that the set of worlds and the accessibility relation can be motivated.

Can we characterize the set of worlds $W$ in any reasonable way? What accessibility relation should be used and why? Within philosophy these questions have been considered in a number of different ways depending on which type of possibility that is under consideration. One can discuss physically possible worlds, logically possible worlds, metaphysically possible worlds, epistemically possible worlds etc. Depending on which kind of necessity that is to be described one can try to specify more precisely the set of worlds and the accessibility relation. Some general results can be derived with only mild assumptions concerning the set of worlds and the accessibility relation.

Another question, that can be asked, is what ontological status the possible worlds have. This question is not settled even when the question concerning the cardinality and structure of the set of worlds has been answered. As stated in the introduction most philosophers do not think of possible worlds, other than our own, as really existing in the same way as our own. For instance Kripke did not consider other possible worlds as existing like ours does. He saw it rather as a way of talking about counterfactual situations or alternative ways the world could be, still his own view on possible worlds was such that he did draw quite a few semantic and metaphysical conclusions from it.\footnote{For a description on Kripke’s views on possible worlds and metaphysical and semantic questions concerning necessity and possibility see (Kripke 1981). The metaphysical ideas of Kripke expressed there are independent from the more formal results of his youth.}

Lewis on the other hand claimed, as has already been mentioned, that other possible worlds are just as real as our actual one. The view advocated by Lewis, namely that all the worlds have the same ontological status, is called
‘modal realism’. Here the differences between various non-Lewisian views on possible worlds need not be discussed, there is no need here to dwell on the finer distinctions and metaphysical assumptions between the different views.

For the purposes of this paper, the relevant difference is just between the view assuming that the possible worlds exist just as our own and other views that do not. If the other views consider possible worlds to be abstract objects, fictions or something else does not really matter. The important thing is that they do not think of possible worlds as physically real and existing concretely in the same way as our own world. Lewis’ main reason for accepting modal realism is that it simplifies systematic work in modal logic, he writes,

I begin in the first chapter by reviewing the many ways in which systematic philosophy goes more easily if we may presuppose modal realism in our analyses. I take this to be a good reason to think modal realism is true, just as the utility of set theory in mathematics is a good reason to believe that there are sets. (Lewis 1986, p. vii)

As most other philosophers, I do not find Lewis’ reasons for accepting modal realism very convincing. On the other hand Lewis has carefully described and elucidated the consequences of accepting other worlds than our own as real. His ideas might thus be useful even if we decide to accept the existence of other worlds for reasons very different from the ones that were of importance for Lewis. One should however bear in mind that to do this, the scheme of Lewis needs to be modified quite a bit. What are called ‘worlds’ in the sense of the physicist are not really worlds in the sense of Lewis. So it is not straightforward to just apply Lewis’ account to many worlds based on considerations in physics. Also since Lewis does not introduce any other restrictions than purely logical consistency he accepts the real existence of many worlds that would not be considered, even in the wildest scenarios suggested the physicists.

Can physics be used to decide W and R, at least for discussions concerning physical possibility and necessity? This does not seem unreasonable. In a sense a theory allows for many solutions describing different scenarios. These scenarios can be identified with the worlds. This will be described further below, where a terminology is suggested, which makes the comparison between possible worlds in the sense of philosophers and the theoretical considerations of the physicists easier.

10 For a discussion on alternative views on possible worlds see (Lewis 1986) where Lewis does not only develop his own view but also criticizes other views which look upon possible world as abstract objects or as linguistic constructions. See also (Divers 2002) for a discussion on various views on the ontological status of possible worlds.

11 An exception might be the Level IV multiverse of Tegmark which will be discussed below.
8.4 Frameworks, theories and models

In this section I introduce a useful terminology. The words in question are 'framework', 'theory' and 'model'. Of course these words are used by many authors. And they can mean different things in different contexts and according to different authors. The words will be used in this article as follows:

**Framework**
A framework is the most general level of theoretical description. It involves a number of central concepts and principles. Within a framework a number of different theories can be formulated. For instance within the framework we can choose parameters and other things so that we describe different theories. A case in point would be the general formalisms describing quantum field theories. This framework allows for various gauge groups and various kinds of particles with different parameters for masses, charges and so forth. So a framework could be said to define a number of different theories.

**Theory**
We get a specific theory by fixing free parameters and other options allowed by the more general framework. For instance quantum electrodynamics, or QED, is a theory in the framework of quantum field theory.

**Model**
A model is what we get when we use the theory to describe a specific scenario. We introduce initial values and/or boundary values and use the theory to find out what happens. A model used in the sense here is a specific solution to a theory. It is not assumed that a model must be uniquely derived from a theory with specific initial and/or boundary conditions. In an indeterministic theory many models can be allowed by the same initial and boundary conditions. Also to calculate a specific scenario it is often the case that a number of simplifying assumptions, approximations and idealizations are used.

I am not under any illusion that these distinctions are very precise or sharp. They are however clear enough for the purposes of my argumentation. The number of levels is to a certain extent arbitrary and more levels could in principle be introduced. The point is that when we go from framework to theory to model, we increase the detailed description so that in the case of a model we have just one specific scenario.

Here follows a number of further clarifying points. Note that the use here differs from the use in many other contexts. According to the terminology here, the standard model of particle physics is not a model but a theory. String theory is better thought of as a framework than a theory, however the lack of free parameters in string theory makes the distinction between what to count as a framework and a theory more blurry, this issue will be addressed and discussed later on.
The use of ‘model’ here differs from the one where people say ‘a theory is a model of reality’. When this is said it typically means that the theory gives a simplified description of reality. Using the terminology above it is, in practice, a model of the theory that is a model of reality, where the second occurrence of ‘model’ means ‘simplified description of’. The expression ‘in practice’ in the previous sentence is deliberate, it is not assumed that a model has to be a simplified description, if the theory happens to be a completely correct fundamental theory and the model for the specific solution of the theory does not involve any simplifications, approximations or idealizations then the description would not be simplified.

If the model of the theory does not involve any simplifications and describes not a part of a world but the whole world we can call this a world according to the theory, if the theory itself is completely correct then the set of worlds of this kind is the physically possible worlds. If the theory is not exactly correct it is just the possible worlds according to the theory.

One needs to distinguish on the one hand, the theoretical description or representation, and on the other what it is supposed to describe. Sometimes people, including myself, can be quite sloppy with this. According to what has been stated above one can define a number of possible worlds according to a theory. The precision with which this set of worlds can be given depends on how precisely the theory is formulated. The set of worlds, possible according to a physical theory, is I would say often more precisely defined than the sets of worlds discussed in many discussions in philosophy. However the worlds, in this sense, are properly better called ‘world descriptions’ they describe a possible scenario. If that scenario does not happen there is nothing physical for which it is a description. We could alternatively see the worlds as abstract objects or fictions but even if one of them correctly describes our own world, or another which happens to really exist, it is then not to be identified by that world but it is to be seen as an abstract or fictitious counterpart of the world.

Take Newtonian physics as an example. I would consider this a theory not a framework; alternative theories with other values of constants or slightly modified laws similar enough to Newtonian physics could be formulated and they would all belong to the same framework. Newtonian worlds are defined precisely enough by the theory. Students learn to explore the set of Newtonian worlds, or rather small parts of Newtonian worlds, when they learn classical mechanics and start to solve problems. Since we do not believe our world to be exactly Newtonian we will never find an exact description of our world among the Newtonian worlds. But there are models of parts of our world that can approximately be described by a Newtonian world or part thereof.

We can in the same way talk of possible worlds according to a framework. If we consider different theories to correspond to different notions of possibility,

---

12 The exception are when the worlds are treated purely mathematically and only the set theoretical details are considered.
we can restrict the accessibility relation, to be such that only worlds that are possible according to the same theory in the framework are accessible to each other.

We do not seem to be committed to the physical existence of worlds existing in the same way as our own just by assuming that a theory allows for many possible worlds. The possible worlds according to the theory are a set of descriptions, whether something physically exists that fits a specific description is another thing. The standard default position among physicists would be to not expect, different worlds possible according to a theory, to exist physically.

The worlds that are here talked about as possible according to a theory are given just in terms of a set of assumptions. The possibility that is needed for the actual practice of calculations and work with a specific theory is only based on these assumptions, it is de dicto. The actual practice of physicists does not require a more metaphysical understanding of possibility. This is not to say that, no metaphysical questions concerning the nature of physical laws can be raised. What is written here is about how we deal with theories. When we talk about real physical possibility we mean according to all correct theories of nature. When using present theories we hope they are in some sense close enough to the real thing.

8.5 String theory and the Landscape

String theorists want to provide us with a satisfying theory for quantum gravity and unify all the fundamental forces in physics within one overarching framework. In earlier days hope was widespread that string theory would be able to give a unique prediction on how our world must be, thus explaining why the constants of nature have the values they have in our world and why exactly the particles we see are the ones that exist. Thus the abundance of parameters of the standard model would be explained.\textsuperscript{13}

Now string theorists have arrived at a different conclusion. The new view is that there are, an immense number of, different sufficiently stable solutions within string theory.\textsuperscript{14} Each such solution describes gives rise to an effective theory with its own ‘constants’, laws of nature and particle content. It is, among other things, the different ways that extra dimensions can be compactified that decide how the laws of that particular world would be. Since there are no free parameters to adjust in the framework of string theory to decide a specific theory the distinction between framework and theory gets blurred. According to the framework of string theory it was traditionally thought that

\textsuperscript{13}Given the formalism in the previous section this would mean that the framework only allowed for one or at least just a few theories. I do not think that even the most optimistic string theorists thought that there would only be one exact model/world with all the details concerning every detail to be found.

\textsuperscript{14}See (Kachru et al. 2003), (Susskind 2007) and (Schellekens 2008).
we could formulate a number of different theories, in a sense close enough to
the one used in this text. As has been told in earlier chapters, we can formulate
bosonic string theory and five different versions of superstring theory, namely:
type I, type IIA, type IIB, heterotic SO(32) and heterotic $E_8 \times E_8$. Now the
different supersymmetric string theories are thought to be related by dualities
which suggests that we should not see them as different theories.\textsuperscript{15}

One might question the use of the word ‘law’ above. It could be argued
that if we really have different solutions with different ‘laws’ they are not
properly speaking fundamental laws and the the word ‘law’ should then be
reserved for the general principles on the more fundamental level. Here I
will nonetheless use the word ‘law’ for what would effectively be laws in the
specific worlds. This would also better confirm to what we have previously
thought to be laws. When needed we can distinguish between fundamental
laws of nature and effective laws of nature. This is however a terminological
question and it is hoped that it will not cause further confusion. This set of
solutions to string theory is the Landscape of string theory. The number of
sufficiently stable vacua in string theory is extremely large, a number often
mention in the literature is $10^{500}$, but this number is a rather rough estimate
and should be taken with a pinch of salt.

Critics have said that such a theory will not be able to predict anything.\textsuperscript{16}
They see this as a failure of string theory and something that would make string
theory basically unfalsifiable. A ‘theory of anything’ rather than a theory of
‘everything’. Is such a critique justified?

In the discussion on the Landscape, versions of the controversial anthropic
principle have been used. Within the string theory community itself there are
also different views on the anthropic argument and some regard the use of an-
thropic arguments to be unscientific. What the anthropic argument amounts to
and the controversy surrounding it will be discussed later on in this text. Here
it suffices to say that the defenders of string theory, who use the anthropic prin-
ciple, think that we should abandon the search for a fundamental derivation of
the specific values of physical parameters in our universe. Instead they claim
that not only one, but a huge number of different solutions are physically real-
ized. That is, they advocate that we should believe in a vast number of parallel
worlds. So this is why string theory has inspired yet another scenario where
we need to consider the existence of a multiplicity of worlds.

It is important to distinguish on the one hand between the Landscape of
string theory in terms of the mathematically possible vacua in string theory
and the further question whether there are physical realizations of these vacua.
Susskind himself distinguish between the Landscape as the purely mathemat-
ical structure and uses the term ‘megaverse’ for the set of physically realized
vacua; a more commonly used term is ‘multiverse’. He writes,

\textsuperscript{15}See chapter 6. for more details.
\textsuperscript{16}See for instance (Smolin 2006b, Smolin 2007) and (Woit 2006).
The two concepts – Landscape and megaverse – should not be confused. The Landscape is not a real place. Think of it as a list of all the designs of hypothetical universes. Each valley represents one such design... The megaverse, by contrast, is quite real. The pocket universes that fill it are actual existing places, not hypothetical possibilities.
(Susskind 2005, p. 381)

This shows that Susskind acknowledges the difference between, what in this article has been called possibility according to a theory and the physical realization of possibilities. What is needed for string theorists is to argue for a plausible mechanism to populate the Landscape with a number of physically realized worlds. The mechanism most popular is eternal inflation. The empirical evidence for inflation is good, but the exact nature behind inflation is not yet understood. Among the proposed solutions many suggest that inflation will go on for ever, parts of the inflating totality will stop inflating thus producing pocket or bubble universes.\(^{17}\) Another candidate mechanism for populating the Landscape is the many worlds interpretation of quantum mechanics.

8.6 The many worlds interpretation of quantum mechanics

The traditional way of understanding quantum mechanics, along the lines of the Copenhagen interpretation, introduces a quite arbitrary division between the observer and the observed world. The Schrödinger equation in itself gives rise to a deterministic and unitary evolution of the wave function or state of the quantum system. When a measurement is made the wave is thought to collapse so that the wave function ends up in an eigenstate to the operator representing the observable that one measures.

How should we understand the collapse of the wave function? What is the ontological status of the wave function and what constitutes a measurement? These are central questions within the foundations of quantum mechanics. Different interpretations of quantum mechanics suggest different answers to these questions.

Most physicists have taken a rather pragmatic approach and decided not to ponder much upon the question of interpreting quantum mechanics. For a long time, only a few physicists and philosophers of physics have worked on this problem. Nowadays, however, I think more and more physicists find it important to deal with this question. An interpretations that has become increasingly popular is the many worlds interpretation of quantum mechanics, henceforth abbreviated MWI.\(^ {18}\)

\(^{17}\)See fo instance (Guth 2007), (Linde 2007).

\(^{18}\)One should probably rather talk of many different many worlds interpretations of quantum mechanics. Since there are quite a few different accounts. Here I disregard this complication.
Everett (1951) wrote the original paper that started the development of the MWI. According to MWI there really is no collapse of the wave function, it continues to develop deterministically and unitarily. The different outcomes of the measurement all happen and are observed in different branches of the wave function. In this way the observer splits and observes different outcomes. How to understand this proposal is however not straightforward. Healey (1984) argues that the interpretation itself is in need of interpreting.

One benefit of the MWI is that it avoids the arbitrary split between observer and observed system. It relies only on the standard unitary evolution of the theory. It is important to note that the MWI is properly speaking deterministic. The appearance of indeterminism only comes because observations happen within one branch.

The MWI is quite weird, from the point of view of common sense, whether that is of any relevance for its plausibility is however not obvious. An internal problem for the MWI is the need of a preferred basis. When should we consider different results to occur in different branches and in which sense does a split happen? The problem is that, while a state might be in a superposition with respect to one observable, it is not in a superposition with respect to another. This is now supposed to be solved in terms of decoherence.

The idea behind decoherence is that the quantum system we study also interacts with an environment that we have not included in the theoretical description. When we trace over the degrees of freedom that we have not included the different possible measurement outcomes in a specific preferred basis will start to behave independently of each other and suppress interference between eigenstates thus forming new separate branches. Decoherence may explain why we do not see interference effects for macroscopic events. It is also supposed to explain what the preferred basis is, which is used for the purpose of describing the branches of the MWI.

Another problem with MWI is the recovery of the Born rules. If different results really happen in different branches in which way is one more probable than the other? While much progress has been made in clarifying and working out details concerning the MWI. It is still just one of many (epistemically) possible interpretations of quantum mechanics. It cannot be seen as part of established science.  

8.7 Tegmark’s taxonomy of multiverses

Tegmark introduces a terminology to discuss multiverses. And I would like to compare this with the discussion that we have pursued in this text. 

---

19 For up to date discussions on interpretations of quantum mechanics, see (Dickson 2007) and (Wallace 2008).

20 The taxonomy of multiverses is presented in (Tegmark 2003, Tegmark 2007, Tegmark 2008), the basic idea is the same but the precise descriptions differs slightly.
fore presenting Tegmark’s description of various types of multiverses, I will explain a little curiosity. It is a not uncommon fantasy, among physicists, to have a T-shirt on which the equations that describe the fundamental laws of physics are printed. For this to be possible the equations must of course first be found. Here is how Tegmark describes the different levels:

- **Level I** A generic prediction of cosmological inflation is an infinite ‘ergodic’ space, which contains Hubble volumes realizing all initial conditions – including one with an identical copy of you about $10^{1029}$ m away.

- **Level II** Given the fundamental laws of physics that physicists one day hope to capture with equations on a T-shirt, different regions of space can exhibit different effective laws of physics (physical constants, dimensionality, particle content, etc.) corresponding to different local minima in a landscape of possibilities.

- **Level III** In unitary quantum mechanics, other branches of the wavefunction add nothing qualitatively new, which is ironic given that this level has historically been the most controversial.

- **Level IV** Other mathematical structures give different fundamental equations of physics for that T-shirt.

(Tegmark 2007, pp. 99-100)

A number of comments can be given to this taxonomy. First when comparing this taxonomy, with the discussion on possible worlds according to Lewis it is only at **Level IV** where we would get something very similar.

At **Level I** we have repetitions within one infinite world or universe. Even if parts of the world can be beyond the horizon from our point of view and thus we will be unable to ever see them they are still just parts of one big world. In a loose sense, one can talk about ‘different effective worlds’ as being part of this bigger world. It should be noted that while different parts of the inflating spacetime may be beyond the horizon to each other, there is not a natural way in which we can partition the spacetime into different effective worlds so that a specific point is in one world only and points in two worlds would always be beyond the horizon to each other. At **Level II** the situation is the same if the different parts are within one and the same space, as indicated by the quotation. Also if the different parts are physically connected in any way, for instance by sharing a history, then I would say that they are better described as being parts of one world.

Tegmark points out that no qualitatively new scenarios are introduced at **Level III**, which basically describes MWI. Because of this, he does not think that MWI should be as controversial as it has been. Here I must say that even if it happens to be the case that no new qualitatively different observable scenarios are introduced there is an important philosophical and metaphysical difference between the scenarios. To deny this, based on instrumentalistic
or positivistic reasons, is not an option consistent with Tegmark’s quite meta-
physically liberal attitude in his approach to different multiverse scenarios.

In the case where there exists a qualitatively identical doppelganger to a
person far away in space; then while being very similar it seems quite natu-
ral to say that they are not the same person. In MWI we can not as easily
think in this fashion. Then the person splits into two, or more, when branches
occur, this is a significant difference. Note that Lewis, who accepts the real
existence of parallel worlds, still did not want to understand them in terms
of branching as is done MWI. For an argument against branching worlds see
(Lewis 2004). There seem to me that there is a real difference between having
two qualitatively similar individuals far away from each other, and the branch-
ing phenomena that occur according to MWI.

This discussion could perhaps fruitfully be connected to all the literature
on personal identity written in philosophy. See in particular (Parfit 1971).
While I think this comparison may be useful I do find the discussion on what
constitutes a person among philosophers to be flawed if understood as a search
for a real metaphysical definition of what a constitutes a person, instead of just
a search for developing a suitable terminology. If MWI is correct our standard
everyday usage would be useless, at a more fundamental physical level and
fail to describe what goes on. Parfit’s view that survival is more relevant than
identity could however then be useful.

If the multiverse, or megaverse, of the Landscape, populated due to eternal
inflation, and the MWI are two descriptions of the same thing, as suggested
for instance in (Susskind 2005), (Aguirre & Tegmark 2011) and (Bousso &
Susskind 2011), then there needs to be quite a lot of conceptual clarification
so as to eradicate the apparent difference. This would, as far as I can see,
involve some form of reinterpretation of how to understand the many world
interpretation and/or the multiverse of the Landscape populated by eternal in-
fation.21 I will not make an attempt to evaluate this suggestion here, however
given that both MWI and the multiverse due to eternal inflation are not well
established parts of science the proposal must at present be judged as highly
speculative.

Of Tegmarks different multiverses, it is **Level IV** that is the most widely
speculative and metaphysical of them all. Here his flights of fancy, in terms of
accepting other worlds as real, can only be rivaled by Lewis. His argument for
accepting different mathematical structures, as all being physically realized, is
reminiscent of how Lewis argues. Tegmark even refers to Lewis in his texts.
While there are some similarities between Lewis and Tegmark, there are also
some differences. Lewis is not sure if he wants to identify the mathematical
with the physical. This is what Tegmark argues would be the case if it was

---

21To reconcile these apparently totally different descriptions. Something at least as drastic as
what was done with the interpretations of dualities, as being descriptions of the same reality, in
chapters 6-7 needs to be done. This could presumably further support my arguments against a
literal understanding of theories.
true that there exists a Level IV-multiverse. Lewis has for instance written the following,

I do not, of course, claim that these complicated mathematical entities are the possible worlds. I cannot believe (though I do not know why not) that our own world is a purely mathematical entity. Since I do not believe that other worlds are different in kind from ours, I do not believe that they are either. What is interesting is not the reduction of worlds to mathematical entities, but rather the claim that the possible worlds stand in certain one-to-one correspondence with certain mathematical entities.

(Lewis 1973, p. 90)

So it seems that Tegmark goes even further in terms of metaphysical extravagance than Lewis. On the other hand, while Tegmark suggests the possibility of the Level IV-multiverse, it is not clear how strong his belief in the existence of this kind of multiverse is. Given his writings I do not think that he really argues for the existence of a Level IV-multiverse in the way that Lewis argues for modal realism.

8.8 The anthropic principle

The use, of different versions of the anthropic principle, is quite controversial. The anthropic principle comes in a number of different versions. The common idea is to use our existence as conscious beings to in one way or another explain why the universe has the properties it has.

The only kind of anthropic principles that will be considered here are versions of the Weak Anthropic Principle (WAP). This is how Barrow and Tipler formulates this principle,

Weak Anthropic Principle (WAP): The observed values of all physical and cosmological quantities are not equally probable but they take on values restricted by the requirement that there exists sites where carbon-based life can evolve and by the requirement that the Universe be old enough for it to have already done so.

(Barrow & Tipler 1986, p. 16)

This principle has been called empty and tautologous. Roughly it states that we can only observe a world which allows for our existence. This seems obviously true and should hardly be thought of as controversial. Some stronger versions of the anthropic principle suggest that the universe must be somehow designed to allow for life or consciousness, this is certainly much more controversial. The weak anthropic principle, being quite empty, is neutral with respect to what further metaphysical conclusions to draw.

---

22For an extensive discussion of the anthropic principle and different versions of it see (Barrow & Tipler 1986).
Our observable world of course allows for our existence but if certain parameters were different we could not exist. See (Barrow & Tipler 1986) for a large number of examples. Thus, it is said, that our universe is fine-tuned. Some see this fine tuning as evidence for the existence of a deity who purposefully has designed the universe we exist in.

An alternative view is to argue that there are many worlds with different values on the fine-tuned quantities. Here I include both genuinely separate worlds and worlds in the looser sense where they can be seen as just parts of one bigger world. This idea is independent of string theory, it is more general. This strategy could also be used by other competing frameworks or theories.

For this to work the framework or theory must allow for many different values of the quantities, some which allow for our existence. There must also exist some mechanism that generate physical realizations of many solutions.

This shows how questions of a theological nature appear when the anthropic principle is discussed. However my view is that theological considerations are better kept away from the scientific discussion. The reason for this is that the theistic view is too unspecific and imprecise. Assuming a deity as an explanation is not precise enough to constitute part of a scientific theory, it is such a vague idea so that it can be adjusted to almost anything. For instance (Page 2007) and (Collins 2007) both think that accepting multiple worlds or a multiverse is compatible with a theistic viewpoint, even though a multiverse scenario often is used to argue against a theistic view of the world. My point is not that I find the claims of Page and Collins to be convincing or important, it is rather that, it shows how vacuous the theistic view is when suggested as a scientific theory.

A story advocating the existence of a multiverse, is by some presented as to to take away the need for a theistic explanation of the universe, see especially (Susskind 2005). However even if that might be so and the very idea of an alternative scenario without deities might be comforting for the atheist, the question is, if this alternative can be formulated in terms that would make it more scientifically respectable, than the theistic alternative. Is any specific of many different alternative multiverse explanations better? Well, without a very precise formulation of a specific multiverse it is, in terms of being science, just as empty as the theistic explanation.

When string theorists argue that we should believe in a multiverse, they often make a comparison with our earlier convictions that we humans were somehow in a central or special position. Then with Copernicus the view that we are in the center of the universe was shattered. The idea is that the string theory Landscape is supposed to just take this a step further. We are supposed

\[23\] In (Stenger 2011) it is argued that the degree to which the universe is fine-tuned for life has been much exaggerated. I will here not try to evaluate his arguments whether or not his conclusions are correct they would not strongly influence the conclusions in this article.
to abandon the in their views naïve idea that our world is the only one that exists.

I do not find that there is anything inherently absurd or implausible with the idea that there are alternative parallel universes, in one sense or another. It might very well be a natural and correct further step in the ‘Copernican direction’. The problem is to find good scientific evidence for this. Even if it is accepted that we have some evidence in favour of inflation giving rise to a multiverse, we do not yet have empirical evidence for a specific theory with multiple possible vacua. It is probably quite difficult to find such evidence.

The very failure of string theory to provide a unique prediction, and due to this suggest the existence of a multiverse, can of course not be taken as evidence for string theory itself. However the rhetoric of expressed for instance in (Susskind 2005, Susskind 2007) and (Schellekens 2008), is deceptively used so as to turn the non-uniqueness of string theory predictions to be seen as counted in favor of string theory. In (Schellekens 2008) it is argued that the view that our universe is unique is similar to previous misconceptions that our status as humans is special or unique. He compares this with the emperor’s clothes from the classic fairy tale. Skeptics might instead wonder whether the comparison with the imperial garments is not more apt for describing string theory research itself. It seems that they have made a virtue out of a necessity.

While string theory might very well be correct, it needs to be connected with experimental predictions in a successful way to be become an established part of science.

Smolin has argued that the anthropic principle in itself is scientifically empty and that it cannot be used to provide evidence in favour of the multiverse suggested by string theory.24

Smolin argues as follows. There are purported uses of the anthropic principle such as the prediction of a resonance in carbon by Hoyle (1953) and the prediction of a small but non zero value of the cosmological constant by Weinberg (1987).

Smolin (2007) claims that the standard form of purported examples of using anthropic reasoning is as follows,

(i) X is necessary for life.
(ii) X is true in our world
(iii) Using the the laws of physics as presently understood, together with other observed facts Y, we deduce that if X is true of our universe, then so is Z.
(iv) We therefore predict that Z is true

Do we actually use anthropic reasoning here? No we do not, since we do not need to invoke the first premise at all to derive the conclusion. If this description of how anthropic reasoning is used is correct, we surely do not need

---

24 See (Smolin 2006b, Smolin 2007).
to use our own existence as a relevant part of the argument. If is assumed that we live in a multiverse (i) can be replaced by (i’),

(i’) We live in one member of the multiverse in which the laws of physics vary. X is necessary for life, so by a selection effect we must live in a universe in which X is true.

Still this modification does not alter the fact that (i’) is irrelevant for the conclusion.

Basically I agree with Smolin’s argument to the effect that we do not really, in a serious way, need to consider our life as a relevant premise for the deduction to go through in this scheme. This will be further explained when I discuss how to make predictions in a multiverse. However the main idea with the use of anthropic principles is to find explanations not predictions. On the other hand I think that successful predictions are needed to make a theory part of legitimate and established science.

One point, which can be made, is that prior to the discovery of a small cosmological constant, most physicists thought that it was precisely zero due to some theoretical reason which we had not yet figured out. Weinberg did not make this assumption and instead calculated a bound for how large the cosmological constant can be given observations of the nature of our universe and then a small positive constant is empirically discovered, I think that this has some relevance in favor of some kind of anthropic reasoning. Even though it actually does fit into the scheme Smolin describes for how purported anthropic reasoning works.

There clearly is something in Smolin’s argument to the effect that anthropic reasoning if described according to the scheme does not really make substantive use of the premise that we exist. However is this the way to reason in a multiverse?

Here it is important to note that Smolin has no problems with the idea of assuming that there exists a multiverse, it is the use of the anthropic principle that he finds problematic.

When he attacks the multiverse of string theory and attempts to show why it is unfalsifiable he argues that the Landscape predicts a completely random configuration of parameters, thus we can not make predictions. Whatever we observe will be compatible with the multiverse of the Landscape. Is this claim really true? As far as I understand we have at present very little understanding of the exact distribution of different vacua in string theory, so this claim might turn out to be false.

He himself suggests another kind multiverse, based on universes developing from black holes, where there is a strong selection mechanism for certain values such that almost every universe has a certain property. We can thus confidently use this multiverse to make falsifiable predictions.
I find this alternative suggestion to be problematic. I leave aside questions whether the suggested mechanism is plausible or not. The problem, that I see, is that I find the need, that the there must be a certain property which almost every universe should have for us to be able to make predictions in a multiverse, too strong a condition to demand. Also it seems to suggest that favored condition, namely producing the maximal number of black holes, also is strongly correlated with the existence of observers. Assuming this is not the case we would not expect to observe the predicted property.

Once the step has been taken that there really is a multiverse, there is something of relevance in the anthropic principle, but how can we use it to make predictions?

For it to be possible to make predictions from a theory, in which it is assumed that there is a real multiverse of worlds, we need to have a detailed understanding of that multiverse and how the different observable properties is distributed over the set of solutions.

Ideally there will be some sharp predictions which can be found if certain observable properties covary in specific ways. If a not yet measured property $X$ must have a given value according to the theory if a number of other properties have certain values we have a sharp prediction. This is an ideal situation and we might not assume this to be the case. We are far from this situation with string theory. What shall we do if we can only deal with probabilities?

Weinstein (2006) has dealt with this problem. He points out that the weak anthropic principle is ambiguous and can be understood in two different ways.

$\textbf{WAP}_1$: “What we can expect to observe must be restricted by the condition necessary for our presence.”

$\textbf{WAP}_2$: “What we can expect to observe must be restricted by the condition necessary for the presence of observers.”

Weinstein argues for the view that we should use $\textbf{WAP}_1$. I agree with this proposal, even though I would like to state it slightly differently, and take it one step further and this step will shortly be described.

It might seem, that by singling out ourself we think of us as special and that is not in accordance with the ‘Copernican view’ of the multiverse which says that we should not think that we are that special. This is a misunderstanding though.

Weinstein’s point is that while both principles are true, $\textbf{WAP}_1$ is the one to use for predictions.

His idea is basically like this; assuming that our theory claims that there really exists a large number of worlds, if we then are to make a prediction we already know that we are in a world allowing for our presence, to predict a further not yet observed property of our universe we should only consider the set of worlds which are still possible (epistemically) according to our theory.
This is very sensible but I think we should take it one step further.\textsuperscript{25} When we have taken the step to assume a multiverse, as part of our theory, and want to calculate the probability of observing a new value of an observable we should use all the information we have about our universe, whether or not they are necessary for our presence. This is just straightforward use of conditional probability where our life or consciousness does not play a specific part compared to any other properties of the universe we happen to inhabit.

For this to work we need to have a well defined measure for the probabilities and a very clear understanding of the specific multiverse that is suggested.\textsuperscript{26} When this is done the apparent mysterious use of our life and consciousness in anthropic arguments disappear. It is also seen how the unnecessary premise disappears from Smolin’s scheme.

So given any suggested multiverse, this is how to make predictions and calculate probabilities. For this to work and for us to make good predictions, we need to understand the suggested multiverse in some detail. String theory is not in this situation yet but it, or some other multiverse theory, might be so in the future.

8.9 Summary and conclusions

When comparing the different ideas concerning multiple worlds in physics and philosophy we see both similarities and differences. The worlds discussed in physics are not as autonomous as the ones suggested by Lewis, they are perhaps more properly described as parts of a world. The exception would be if the flamboyant metaphysics of Tegmark’s Level IV-multiverse is accepted.

When we use specific theories to talk about possible worlds, the default position is to not consider alternative solutions to be real existing worlds. To take the step and claim that there is a multiverse, some plausible mechanism must be suggested, MWI and eternal inflation are suggested candidates.

To be able to make predictions, based on a multiverse theory, we need a very clear understanding of how the specific multiverse theory distributes observable properties among the universes. Then there is a possibility that falsifiable

\textsuperscript{25}To be fair I think that Weinstein already takes this step in his article. But it is not explicit in the formulation of WAP\textsubscript{1}.

\textsuperscript{26}The general problem of making predictions in a universe has been discussed in (Aguirre 2007). He also discusses the strategy to use everything we observe as input. He however sees some problems with this strategy. It may happen that we have to accept a theory for which the observables we see in our universe is extremely rare. I do not see this as a huge problem, we might very well be atypical, the important question is what the theory predicts given the observations we have already observed. If this prediction fails we have reason to abandon the theory. This is definitely the case if we observe something forbidden according to the theory.

If we often observe something highly unlikely according to the theory we of course have some evidence against the theory, although it is not completely falsified. The new observation will be taken into account, when further observations are done.
predictions can be made. When making predictions we should base them on everything we have observed, not only what is needed for our existence; this makes anthropic arguments less mysterious.

The multiverse of string theory is not yet, and might never be, fully understood so as to make testable predictions. It might turn out that good testable predictions from string theory eventually can be made. If these predictions are then confirmed I think string theory will be generally accepted as part of established science. On the other hand, if it turns out the distribution of solutions in string theory is such that no good predictions can be made, string theory will never reach a truly scientific status. In this case, supporters of string theory may continue to believe in their theory as a matter of faith.
9. Thought experiments and quantum gravity

This chapter is like the previous one also based on a manuscript that is work in progress.\(^1\) Hence the comments made in the beginning of the previous chapter applies to this chapter as well.

9.1 Introduction

In the history of physics thought experiments have been used successfully by many thinkers. Notable masters of the use of thought experiments are Galileo and Einstein. The study of thought experiments in science and philosophy, has in recent years become a topic of increased interest within philosophy of science.\(^2\)

Can a thought experiment, in itself, provide reliable and convincing justification for new beliefs? Some thinkers have denied this, according to them thought experiments are inherently conservative and should not be used in serious science; at least they should not been considered to provide reliable evidence. Others think that they are useful and that they can give us new reliable information. The question is how can we use thought experiments in science; what are their epistemological status?

Since, in quantum gravity, there is a lack of empirical data which can be used to test the various approaches and theories. One might be tempted to look for justification elsewhere. Are thought experiments such that they can provide us with reliable information about what a theory of quantum gravity should look like? There are a few thought experiments to be found in the literature on quantum gravity. One is described in (Eppley & Hannah 1977). This though experiment has been discussed by philosophers of science.\(^3\)

Instead of directly discussing thought experiments in quantum gravity I will argue for a specific view in the debate on how one should understand the epistemic status of thought experiments. Based on my general conclusions an assessment of the possible role of thought experiments in the development of quantum gravity will be given.

Different accounts of thought experiments have been given. In this chapter I will give some background information on the current discussion on thought

\(^1\) (Matsubara work in progress, b).
\(^2\) A few books on this topic are (Horowitz & Massey 1991), (Brown 2011), (Sorensen 2012) and (Frappier, et al. 2012).
\(^3\) See (Callender & Huggett 2001, pp. 5-12) and (Shumelda 2012).
experiments. First a general rough description of how a thought experiment is supposed to work will be given. This is followed by a short presentation of a taxonomy of thought experiments taken from (Brown 2011). After that I will briefly present two views on thought experiments, namely Brown’s platonism and Norton’s view where thought experiments are claimed to be understood in terms of ordinary arguments. I will follow this by arguing against Norton’s view of thought experiments as arguments. I will present and expand on previous critique by various authors. After that I will share my own views on how to understand thought experiments in terms of an alternative to Brown’s platonism. I will also address how this view informs us on when thought experiments are reliable and when they are not.

Finally, in the end of the chapter, I will discuss how my conclusions regarding how thought experiments work informs my view on thought experiments in quantum gravity.

9.2 The anatomy of a thought experiment

How are we to describe or define a thought experiment? I do not think that a general definition, covering all kinds of thought experiments, is needed or even desirable as a starting point. There are many different thought experiments that work in different ways.  

So it should clearly be stated that there is no need to find a single account of how all thought experiments are supposed to work. Thought experiments are best discussed by giving a number of explications of different kinds of thought experiments. Their epistemological status, or the way they are to be understood, need not be the same, for instance we might say that some thought experiments can without serious loss be understood in terms of arguments while others may not.

In this section I just want to give a very rough general description, adapted from (Brown 2011, p. 51), of the anatomy of a thought experiment. Schematically we can understand a thought experiment like this:

Theory & Background → Phenomena → Result.

This schema is applicable to both ordinary experiments and thought experiments. In the first step the experiment, whether real or imaginary, is set up. This is part of the background in the schema, we also assume previously accepted theories and other not necessarily articulated knowledge to be captured in the background. If it is a real experiment something will happen, there will be real phenomena which we can observe. In a thought experiment what will happen takes place in the mind, we anticipate a certain result, in one way or

4On this I agree with what is written in(Brown 2011, p. 32)
another. What we in this case imagine are the phenomena of the thought experiment. Finally we interpret the phenomena and our conclusions are the results of our efforts. Below I will repeatedly write that conclusions follow from phenomena. This is meant in the above described way. The word "conclusions", used in this sense, is meant to be more general than a more traditional use, where a conclusion is the result of a derivation from a number of premises.

9.3 A taxonomy of thought experiments

In (Brown 2011), a taxonomy of thought experiments is given. It is not supposed to be exhaustive and also it is not claimed that it is the only reasonable way of classifying thought experiments. Here the taxonomy will be used without discussing alternative possibilities.

9.3.1 Destructive thought experiments

The destructive thought experiments are used to undermine the claims of a theory, view or opinion. It is similar to a reductio ad absurdum but using a thought experiment. This works by showing that the previously held theory leads to problems, either internal inconsistency or conflict with other strongly held beliefs. This would give you reason to make revisions in your total view of the world. That is of course that you accept the result of the thought experiment.

An interesting specific kinds of destructive thought experiment can be distinguished, namely the counter thought experiment.

Counter thought experiments

This is a special kind of destructive thought experiment in which the background assumptions from another thought experiment are accepted. The aim of the counter thought experiment is to undermine the conclusion of the other thought experiment. It does not question that the conclusion follows from the phenomena. Instead it questions the step from theory and background to the phenomena in the first place.

Basically an alternative story of what would happen is given. This does not conclusively rule out that the original conclusion is the right one, but it can undermine our trust in that conclusion. A famous example of this is the way in which Galileo countered the Aristotelian argument against the theory that the earth moves around the sun. The Aristotelians claimed that if a stone was dropped from a tower on a moving earth the stone would not fall to the bottom of the tower, instead it would fall a bit away from the tower since the earth moves. Aristotle countered this by arguing that even if the earth moved this was not what was going to happen, the details of his new thought experiment
involved considering how things would move inside of a cabin of a ship that
smoothly moved along a water surface.\(^5\)

An interesting question to think about is how often counter thought experi-
ments can be made. Is it possible, that in many cases, we just have not found
a good counter experiment to another thought experiment? If so that would
clearly undermine the reliability of thought experiments.

9.3.2 Constructive thought experiments
The idea of a constructive thought experiment is not to refute a theory, instead
they are used to support a theory. The constructive thought experiments can be
further subdivided into, conjectural, mediative and direct thought experiments.

**Conjectural thought experiments**
A conjectural thought experiment is used to find a new theory. The thought ex-
periment is used to find phenomena, understood in the way described above..
When the phenomena are accepted a conjectural explanation can be given for
why this would happen. The last step I would describe basically as abduction
or inference to the best explanation based on a thought experiment.

**Mediative thought experiments**
Used to facilitate conclusions from specific well articulated theory. That is we
use a theory that we already have faith in. The thought experiment is used to
help us draw certain consequences of that theory.

**Direct thought experiment**
Here one is not supposed to start with a well developed theory. Instead one
starts with a strongly convincing and unproblematic phenomenon based on a
thought experiment and as a consequence of these one is able to formulate a
new theory. In (Brown 2011) it is stated that the borderline between mediative
and direct thought experiments is blurry, since we always start with some as-
sumptions. The blurriness is to be explained by the fact that there is no sharp
boundary between well-articulated theories and general beliefs.

The difference between a direct and a conjectural thought experiment is, as
far as I can see it this; the ‘logical’ order between the phenomena established
by the thought experiment. In a direct thought experiment the order is like
this,

\[
\text{Phenomena} \rightarrow \text{New theory},
\]

while in a conjectural thought experiment it is like this,

\[5\text{See (Brown 2011, p. 34) for a description of this. For a translated version of the original account, see (Galileo 1967).}\]
Above I have used scare quotes on ‘logical’ since it does not have to be a strict logical relation. It can certainly be something weaker, so the arrows above are not to be understood in terms of logical implication.

9.3.3 Platonic thought experiments

These thought experiments are both destructive and constructive at the same time. They refute one theory and establish a new one. It is particularly for this kind of thought experiment that Brown argues that we need an ability to “see into Plato’s heaven”, hence the name. Here I will just use it as a label for a specific kind of thought experiment. I do not mean that by calling a thought experiment ‘platonic’ it is assumed that a platonist epistemology is accepted. Presumably a critic of Brown’s platonist explanation, such as myself, would prefer a different more neutral term. I choose however to not suggest a new term in this context.

Brown does not argue that the method use thought experiments is infallible. This holds also for those that fall in the category of platonic thought experiments. This is clearly admitted by Brown, he says that our ability to see the laws of nature in the platonic way is fallible. But what is meant by this supposed ability to see the laws of nature and why do Brown argue for the need to explain at least some thought experiments with the help of this ability? It is now time to give a short presentation of Brown’s views on this issue.

9.4 Brown’s platonism

In the philosophy of mathematics Brown defended a platonist view before arguing that a platonist view could also contribute to our understanding of how we discover the laws of nature. According to (Brown 1991, Brown 2011) we have a certain ability to ‘see’ the laws of nature a priori. Thought experiments are the vehicle which we use to thus discover laws of nature. This is supposed to work in a way similar to how we understand mathematics according to a platonist view on the foundations of mathematics.

All this seems quite strange for a person, such as myself, with strong empiricist leanings. Later on I will present an alternative view on how we are to understand the instances where Brown resorts to the platonist explanations to describe how a thought experiment works. Before that I want to describe Brown’s view in a bit more detail.

Why does Brown claim that we can reach conclusions from thought experiments, in an a priori way? The reason why thought experiment are supposed to be a way of reaching new knowledge a priori is, that we use no new data, the
conclusion is not a logical truth and it is not derived from previously accepted empirical truths.

Brown does not claim that we need to consider this ability to see the laws of nature in Plato’s heaven to account for all kinds of thought experiments. It is primarily for the platonist thought experiments that we need to rely on this ability. Also it must be noted that while it is supposed that we can use this ability to come to conclusions a priori, it is explicitly stated that we are fallible when doing so.

Our fallibility is not thought to be a serious problem, since all agree that the ordinary senses also are fallible, and yet this does not mean that we rule out empiricist methods based on the ordinary senses. Our ability or sense which helps us to grasp the nature of abstract objects is not different. That we have no idea how this weird sense is supposed to work is also countered by arguing that we know very little in detail about how our ordinary senses work.

Not much can be said about how this kind of a priori access is supposed to work except the following. It is a fallible ability which is also supposed to be non-linguistic and not rely on a Kantian account of how our perceptions work. I am quite sympathetic to the stressing of the non-linguistic as will be made clear later on. On the other hand my own view on the issue has a slightly Kantian flavor since it partly refers to some hardwired ways of conceptualizing our perception. It is however a fallible version where I rely on evolution for the explanation of how our perceptions work.

Brown uses platonism regarding mathematics, as a starting point to support his platonist views concerning thought experiments and the laws of nature. The laws of nature are understood by Brown in terms of necessary relations between universals. This view of natural laws has previously been advocated by Armstrong, Dretske and Tooley. Universals are supposed to be abstract objects, that are existing, for real, independently of us. This is, after all, just the way numbers and other mathematical objects are understood according to a platonist view of mathematics. Hence if we can understand some abstract objects in an a priori fashion why not others.

The arguments, given in favor of platonism, have very little force if one does not already accept a platonist view of mathematics and the Armstrong-Dretske-Tooley account of natural laws.

I will not present further arguments, put forth by Brown, in defense of this platonist view concerning our access to the laws of nature via thought experiments. Instead I will present Norton’s alternative account of thought experiments.

---

6See (Armstrong 1983), (Dretske 1977) and (Tooley 1977).
9.5 Norton’s view that thought experiments are arguments

According to Norton, thought experiments are basically just a kind of picturesque arguments. My description here is based on (Norton 1991) and the presentation of Norton’s views in (Brown 2011).

Norton wants to give an empiricist account of thought experiments. Norton claims that thought experiments are arguments which:

1. posit hypothetical or counterfactual states of affairs, and
2. invoke particulars irrelevant to the generality of the conclusion.

Note that Norton explicitly says that thought experiments are arguments. His view is that since thought experiments provide or at least purport to provide knowledge of the world without the use of new empirical data, there is only one, non-controversial way of doing this, namely through arguments based on previously justified theories and beliefs. So thought experiments are basically arguments from previously held beliefs about the world.

The introduction of hypothetical or counterfactual, states of affairs is supposed to capture the thought-like aspect of thought experiments.

That the thought experiment, invokes particulars irrelevant to the generality of the conclusion, is supposed to capture the experimental part of the thought experiment. Particulars can be eliminated by being part of a counterexample to general principles. Another way of eliminating the particulars is to use an argument based on an inductive step where the particulars are seen as typical.

It should be noted that Norton allows for all kinds of different arguments. He does not only tolerate strict deductive arguments, inductive arguments are also accepted. Also he allows the introduction of general philosophical principles or assumptions as parts of the premises of the argument. Presumably I assume that if general principles are introduced, for which we do not have good justification, the argument would be considered bad.

Basically Norton says that if a thought experiment is not an argument, the conclusion can not be credible.

Norton advocates the elimination thesis which states the following:7

The Elimination Thesis: Any conclusion reached by a (successful) scientific thought experiment will also be demonstrable by a non-thought-experimental argument.

---

7The specific formulation of the thesis used here if from (Gendler 1998). She refers there to (Norton 1991) but there the formulation is different and also involves the discussion on how to eliminate particulars from the thought experiment. Gendler also refers to another paper by Norton written in 1996, I have however not read that article to check how close Gendler is to Norton’s formulation in that paper. The exact formulation used by Norton is however not needed in this context.
Norton says that this is always, in principle possible, but that it can be very difficult in practice. Thought experiments are used when straightforward arguments are difficult to formulate.

A good thought experiment is or can be reconstructed as a good argument, a bad thought experiment will at best be or be reconstructible as a bad argument.

A distinction which is not so carefully stressed in (Norton 1991) is the distinction between saying that a thought experiment is an argument and the claim that it can be reconstructed as an argument. Norton seems to think that if we can reconstruct a thought experiment to an argument we can basically say that it was an argument all along. This is not agreed upon by many critics of Norton’s view which I will discuss further below.

I think that Norton’s view is too closely connected to a linguistic and statement based view on theories. There are aspects of thought experiments that could not just be seen as being as reducible to arguments. However I do think it is the case that many thought experiment, without losing anything essential, can be reconstructed into ordinary arguments.

9.6 A critique of Norton’s view

Here I present criticism against Norton’s views that thought experiments are arguments. I will refer to a number of earlier criticisms and add in further comments of my own.

In (Bishop 1999), it is argued that thought experiments can not be arguments since a single thought experiment can be reconstructed as more than one argument. This is however not a very serious criticism since it can simply be countered by claiming that Bishop does not have the same understanding of a thought experiment.

If a thought experiment is supposed to be the whole chain from the theory and background where the initial assumptions of the thought experiment is given via the phenomena to the final conclusion. Then the example given by Bishop is just a case where there are two different thought experiments, with the same initial setup. A defender, of the Norton’s argument view, can also say that one argument is a bad argument and the other a good argument since one argument misses relevant theoretical assumptions. Note also that in the taxonomy of thought experiments, a counter thought experiment is considered to be a different thought experiment to the one it counters even though the starting assumptions is supposed to be the same. Given this it follows that I do not accept this as good criticism against Norton’s view, I decided to include it here nevertheless. I now turn to more promising approaches to attack Norton’s views.

In (Gendler 1998), Norton’s account of thought experiments is attacked. An important part of the criticism, is that without the prior use of a thought experiment, certain assumptions needed for turning the thought experiment
into an argument would seem very ad hoc. Just reading the argument afterward would be less convincing than going through the thought experiment so the justificatory force is not based on the reconstructed argument.

Gendler points out that the Elimination Thesis of Norton is ambiguous. So Gendler present two more precise formulations:

1. The Dispensability Thesis: Any good scientific thought experiment can be replaced, without loss of demonstrative force, by a non thought-experimental argument.

2. The Derivativity Thesis: The justificatory force of any good scientific thought experiment can only be explained by the fact that it can be replaced without loss of demonstrative force, by a non-thought-experimental argument.

She attacks the first thesis and show that the reconstructed argument has less justificatory force than the original thought experiment. Of course if one does not accept the first thesis it follows that one can not accept the second thesis.

The example Gendler focuses on is Galileo’s thought experiment with falling bodies. This is used to attack the Aristotelian theory and conclude that (disregarding air resistance) all objects regardless of their mass fall with the same speed.8

If the speed of a heavy object is proportional to its weight, as assumed by the Aristotelians, it seems that when two objects one heavy such as a cannon-ball and a light one such as a musket ball are connected by a piece of string, it follows from the Aristotelian theory that this composite object both fall slower and faster than the heavy object by itself. This is since on the one hand the musket ball would slow down the heavier object. On the other hand the composite system is heavier than the heavy object by itself and should hence fall faster. This contradiction in the Aristotelian system is supposed to destroy our belief in that system. The thought experiment is also supposed to give us the conclusion that all objects (when air resistance is neglected) fall with the same speed.

The Galileo thought experiment is the prime example of the kind of thought experiment that is labeled as platonic according to Brown’s taxonomy. It involves a destructive and a constructive part. Let us first focus on the destructive part.

We certainly do find the destructive part of the argument convincing, however Gendler points out that if it was just a plain argument the Aristotelian could deny a number of hidden assumptions in the thought experiment, to avoid the conclusion.

She points out that the contradiction can be evaded by claiming that when two objects are tied together, they can not automatically be thought of as one object. She formulates a function for how the two masses, that are held together, may be held together with different levels of connectedness $C$, where

---

8See (Brown 2011) and (Galileo 1967) for further details on this thought experiment.
$C$ varies between 0 and 1. For the case where the connectedness is 0 it is correct to describe the situation in terms of the small mass slowing down the bigger and for the case of 1 the mass would fall as one heavy object and hence faster than the heavy body.

This solution does however not at all strike us as plausible. Gendler says that the reason for why these Aristotelian countermeasures seems so ad hoc is that they are in conflict with our tacit knowledge about the physical world. When we think through the thought experiment we do not feel the need to explicitly state all needed assumptions. Also we are probably not aware about what all these assumptions are, the thought experiment may however help us to state these assumptions in a way which could be put into an argument. But they would then seem very much ad hoc. She says in (Gendler 1998, p. 415) that, “The justificatory force of thought experiments is thus parasitic on the extent to which the messy twisted web of background beliefs that underpin our navigation of the world are rightly considered knowledge.” This means that she claims that the reliability of the thought experiment depends on the reliability of this background knowledge. I agree with this and will develop this theme further.

Gendler talks about the constructive and participatory aspect of thought experiments. Gendler describes ‘experiment in thought’ as a participatory part of the experience of a thought experiment. We basically think about the situation and explain what we expect to happen. When doing this we use our previous theoretical beliefs about the world but also an instinctive and unarticulated ability to judge what would happen. She does not talk widely about this aspect, but I think this could be well explained at least in part by the ideas concerning mental models which will be described below.

Gendlers view is in some respects similar to the views in (Kuhn 1977 [1964]). In that text Kuhn expands on the old view that thought experiments can only give us knowledge about our conceptual scheme and not about nature. He thinks this is inadequate since thought experiments show how certain situations can not be handled by the concepts used in the old scheme and this can be used to reconsider our theory and our conceptual scheme. He like Gendler also stresses the importance of tacit knowledge.

Here are two quotations from Kuhn:

If a thought experiment is to be effective, it must, as we have already seen, present a normal situation, that is, a situation which the man who analyzes the experiment feels well equipped by prior experience to handle.  
(Kuhn 1977 [1964], p. 252)

‘I want now to argue that from thought experiments most people learn about their concepts and the world together.’
(Kuhn 1977 [1964], p. 252)
Kuhn is not a realist concerning science and does not assume an independent world, in one sense or another. Hence the difference between learning about scheme and world is blurred. I think this is not much of a problem, we can accept part of Kuhn’s view while not accepting the relativist reading of Kuhn. There is also some controversy on how to understand Kuhn, some criticize the relativist reading of Kuhn. The blurring between learning about the world and learning about concepts can also be found in Quine, who is generally not seen as anti-realist or a relativist concerning science.

A problem with Norton’s view is that it describes scientific theories too much in terms of a set of propositions. It stresses the linguistic aspect of theories and still seems to rely to a large extent on the old fashioned logical positivists views on the logical structure of theories. This view identifies a theory with a set of sentences or propositions abstractly derived from a number of axioms. While this can be a useful idealization to use for some situations, it is a very simplified and even distorted description of how actual physical theories are used.

The argument that theories, in principle could be reformulated like this, is not relevant if we want to discuss actual practice with theories at their present state. I am also not sure to what extent they in principle could be reformulated like that, especially when we discuss developing theories which by their nature contain lots of vagueness and different possible interpretations. Possibly a unique and well formulated strict logical formulation could be given for a theory that has stopped evolving. Such a theory would however most likely be a ‘dead theory’ which we have abandoned and do not believe to be strictly true. Of course this should not be understood as me being against formalism, mathematical or logical, when used in the right context. It is the misguided use of this in a way that idealizes in such a way that it deforms or misrepresents the function of theories that I find problematic. Actual theories in physics do of course contain a lot of mathematical formalism and large parts of the theories are given in a quite axiomatic fashion. But there are other aspects of the theories that depend much more on everyday language, metaphors and analogies.

A lot of the insight that scientists get is of the tacit kind and not explicitly formulated. It is the capabilities, that the actual practice of working with the theories gives rise to. This might be slightly unfair to Norton since he does not demand a strict formalization, yet I think it is a fair point that he overemphasizes the linguistic aspect of our thinking. As has been established above Gendler and Kuhn stresses tacit knowledge and background beliefs which to a certain extent can be understood in terms of a defense of the non linguistic aspects of our thinking. Another kind of argument against Norton tries to

---

9In this paragraph I use the word ‘theory’ in a quite loose sense similar to how the word is used in ordinary practice among scientists. Maybe what I describe as a developing theory should be described as a succession of theories instead, nonetheless the vagueness will still be there even if we focus on a specific moment of this development.
explain the non-linguistic aspect of thought experiments in terms of so called mental models.

So what are mental models? The following discussion of mental models is based on (Nersessian 1992) and (Miščević 1992). Basically mental models are supposed to explain how we can visualize certain situations and how we reason visually.\(^\text{10}\) Miščević gives the example where we in our mind imagine how our room would look if we moved our furniture in a certain way as preparation for our decision to actually move our furniture in a certain way. This very humdrum example is presumably familiar to all. We all have experience of thinking in this visual way and I do not think that ought to be controversial. Also we do not only have to rely purely on introspection for evidence that we use this kind of visual reasoning.

There are experiments that support the view that we use mental models when reasoning. Nersessian argues that we typically do not think by making formal deductions on statements and propositions. The mental models are supposed to work by converting a linguistically presented story to a mental representation. When asked questions, concerning the original story, we of course give answers linguistically, but the evidence seems to support the view that we often reason not by deducing consequences formally from the statements of the original story but by ‘looking’ at our mental models. The evidence from experiment show that when a story is such that a ‘clear picture of the situation can be made in our head’ we can answer specific questions much faster than in other cases where this is not the case, even if the logical difficulty of making the deductions based on the statements in the story is not more difficult.

I think that it is correct that much reasoning of scientists is done in terms of visualizing things and that it is only later that their claims are articulated linguistically. I guess, based primarily on prejudice and personal experience, that typically physicists are very visual in their thinking while philosophers are more linguistically oriented. This I think, if it is true, is part of the explanation for why many accounts of scientific theories by philosophers are so distorted.

That we thus can imagine, visualize and anticipate certain scenarios in terms of mental models is explained by evolution. It is clearly advantageous to be able to use this ability. Nersessian (1992) claims that the part of our brain handling visual stimuli is far more developed and that this ability evolutionary is older than our linguistic ability. Presumably animals also have this ability to form mental models, that would explain how birds can solve completely new puzzles without training. I have seen examples where birds get a nut by pushing levers, pulling out sticks and so forth on an artificial model made

---

\(^{10}\) What about the other senses, can we not have mental models for them too? Well I can in my head certainly conjure up sounds, smells touch and so forth but I have no idea of how I can reason with those internal representations. They can elicit emotions but I can not see a way of reasoning with them. If this is a common experience among humans it is probably a central part of the explanation for why we do not find thought experiments that involve the other senses.
by plastic building blocks where they can see the nut, but not get it without solving a long set of visual problems.

There are some problems with the reliability of using mental models. Sometimes we add in things that are not explicitly stated in the original story. This can be explained by how our other beliefs, theories and prejudices influence things. Also since this ability is supposed to be an evolutionary adaptation it should be much more trustworthy when dealing with everyday situations, compared with modern theories describing very unfamiliar situations.

Concerning the reliability of our mental models Miščević in the end of his paper writes:

Our basic capacities are probably better tailored for moving furniture than for doing quantum mechanics.
(Miščević 1992)

9.7 An alternative to platonism

In this section I will formulate my own view on thought experiments. My own view is heavily dependent on the views of those attacking Norton’s account of thought experiments. I do not claim much originality I rely heavily on the ideas of Gendler and the ideas concerning mental models. I will however try to connect these two accounts. I also want to address the reliability of thought experiments and their epistemological status in different contexts. I also do not want to rely on platonism in my account of thought experiments. The illusion of the a priori vision is explained by our previous experiences and perhaps expectations that have not been articulated of how nature works combined with our ‘hardwired’ ability to form mental models.

In thought experiments we often use our ability to visualize situations using our ability to construct mental models. This ability is part of a kind of everyday, or common sense, understanding of the world which I here denote \( E \). This is of course a very vague description. That it is given a specific notation is not at all to indicate anything more precise, it is for practical purposes only. Besides the innate abilities to form mental models \( E \) is also supposed to include tacit knowledge gained and learned by interactions with the world. This knowledge has not been explicitly formulated into an organized theory and is to a large extent nonlinguistic. All of it does not have to be nonlinguistic, some parts of \( E \) can be explicit but it should not be part of a more sophisticated theory of the world. This \( E \), I assume is to a large extent common among humans. I admit that it is an idealization, of course it will differ between people because we have different experiences. Also I think it is difficult to draw a strict line between what is captured in \( E \) and what is given in a theory since my view on theories is such that theories also depend heavily on less articulated ideas.

125
In $E$, I want to capture both the relevant ideas from mental models and some of the less articulated or tacit aspects of our previous conceptual scheme. This is supposed to include Gendler’s ideas on how thought experiments work.

In a thought experiment an individual confronts his or hers $E$ with a certain described situation and typically also assumptions from another theory $T$. By doing this we bring out hidden assumptions in $E$, this might lead us to reconsider our view on $T$ or suggest new hypotheses. In situations where $E$ is reliable we can then formulate reliable thought experiments, but if $E$ is applied beyond areas in which we have reason to trust it we should be careful.

In some situations we confront two theories $T_1$ and $T_2$ without really letting $E$ influence our evaluation. The more that we only need to rely on explicit statements from the theories the more the thought experiment can be correctly described according to Nortons account. These thought experiments can without loss be reformulated as arguments; in these situation thought experiments can be used to show inconsistencies between assumptions made in the different theories. Often the situation is such that it can be very difficult to decide which assumptions in the theories, that most trustworthy in the situation, thus making it hard to decide what conclusions we are to draw from such inconsistencies.

### 9.8 Implications for thought experiments in quantum gravity

In this chapter I have discussed how we are to understand thought experiments in physics. Thought experiments use a combination of linguistic and non-linguistic background assumptions. Our ability to make mental models also play a very important role. Since our ability to do mental models is explained in evolutionary terms and is adapted for survival in everyday situations, it can be trusted in such situations. It is however much less reliable when we try to deal with situations far removed from ordinary experiences. This is obviously the case when thought experiments are used to guide us in the search for a theory of quantum gravity.

The thought experiments in quantum gravity research that have some chance of being reliable are the ones that more or less explicitly and without loss can be treated as arguments. Then consequences of different assumptions can be investigated, the problem is that even in such situations it can be difficult to come to correct conclusions. This is because we are in a situation where we need to reconsider even well established principles. Different researchers judge the relevance of these principles differently and would come to different conclusions regarding which principle to abandon when complications arise.

The thought experiment in (Eppley & Hannah 1977) is supposed to show that we are forced to quantize the gravitational field, since otherwise we need to accept either superluminal signaling or a violation of energy-momentum
conservation. As is cleverly pointed out in (Shumelda 2012) the conclusion is not watertight. Interestingly, it turns out that it depends on the interpretation of quantum mechanics that is assumed.

The interesting point here is not so much, whether we are forced to quantize the gravitational field. There are strong arguments for this even without considering the thought experiment of Eppley & Hannah. What I find interesting is that it is an example where the interpretations of quantum mechanics have relevance for which conclusions to draw regarding question connected with quantum gravity.

It might be the case that the interpretation of quantum has relevance for the formulation of quantum gravity. These kind of argumentative thought experiment can, as Shumelda has pointed out, be used to give constraints on which combinations of principles concerning GTR and quantum physics, including interpretative issues, that can be used when formulating a quantum theory of gravity.
10. Summary

In this dissertation, a number of different philosophical questions have been discussed, relating to modern research in quantum gravity. On the one hand, there is the unavoidable question concerning the scientific status of these theories. The conclusion drawn here is that research in quantum gravity, as such, can not be seen as unscientific. But it is also fair to say that no theory at present can be seen as part of established science. This is due to the lack of empirical results. All approaches to quantum gravity are very much work in progress, where the scientists are actively engaged in turning their speculations into respectable and empirically successful theories.

Given this I think it is important that the scientists are open with the speculative nature of their work. So they should not take success for granted and avoid presenting their results as if they already were part of established and empirically tested science.

It was also argued that in the case, where there is no real empirical progress, it is especially important to have a pluralistic approach and try many different approaches. As seen from the examples of dualities in string theory, very different descriptions may nevertheless be connected. The possibility that the different approaches can all contribute, by providing different parts of the puzzle, should not be forgotten.

It was also argued that it is important to distinguish between what is mathematics and what is physics in a theory. Given the fact that string theory, and other approaches to quantum gravity, are so far removed from any more concrete empirical data, it is especially difficult to keep the mathematics and physics apart. Already, without the support from considerations based on dualities, a strong case can be made that, we do not really understand the meaning of many terms used in highly advanced theories of physics. The dualities in string theory can be used to strengthen the claim that, we cannot just take a specific theory formulation literally and with good reason believe that, it in all respects describe an independent reality.

I advocate that we should accept some weaker form of verificationist semantics. This view is not compatible with standard scientific realism. However, it is compatible with how physicists themselves view their theories. I also think there are good independent reasons for why this is how we should understand theories.

Still it is equally important to stress, that it is clear that much of the work of physicists is difficult to understand, without attributing a realist view on scientific theories to them. The search for unification and the willingness to
advocate such claims that there are parallel universes or worlds, are signs that indicate that physicists tend to have a realist view of their theories, to some extent at least. They do not view their theories merely as tools for predictions. While this in itself is just a descriptive account of the attitude of physicists, I have argued that there are good reasons for this tension between scientific realism and anti-realism. I claim, that the best way of understanding theories of modern physics, is in terms of some form of intermediate realism. While structural realism, of one brand or another, might be a solution I can not wholeheartedly defend any specific version of structural realism.

I have also discussed the various ideas concerning multiple worlds or universes, which have cropped up in physics. I compared them with the discussion on possible worlds from philosophy. Much of what is claimed in these discussions seems very speculative. There is a rhetoric where, the failure in string theory to predict a specific vacuum, is used to argue that an extreme multitude of different solutions are physically realized. Then the weak anthropic principle is used to explain the fine tuning for life of the universe. For this to work they need to argue for some form of natural mechanism to explain why we have many physical realizations of different ‘universes’. Even if it is agreed that the existence of bubble universes is a reasonably straightforward extrapolation from inflationary physics, I cannot accept the view that this can be turned into an argument in favor of string theory. The reason being, that doing so is just taking for granted that string theory is correct, without providing any new empirical evidence. It might well be the case that there are other theories which also produce a large number of solutions and could also be used in the same way. To make these claims scientifically legitimate, I argue that we need a theory where the distribution of properties among the various solutions are very well understood and where predictions can be made. I argue that when such predictions are made we should condition on all our information, not just such information which is of relevance for the existence of intelligent creatures. This I think demystifies much of the anthropic principle. However, in this context we need to distinguish between questions of prediction and questions of explanation.

Since there is no experimental support, for any of the different approaches to quantum gravity, it might be relevant to ask the question whether thought experiments might in any way replace conventional experiments. In the debate on thought experiments, it has been argued that thought experiments, in themselves, can be used to provide evidence that sometimes is reliable. Given that I explain the cases where thought experiments give us reliable evidence, without being reducible to arguments, mostly in terms of evolutionary adaptations; I find that thought experiments in quantum gravity cannot be very reliable. The concepts needed and the situations described are too far removed from everyday experience, for any evolutionary hardwired intuitions, to be of much help. Thought experiments, which are of the purely argumentative kind can of course still be used to investigate questions of consistency in assumptions and
consequences of various assumptions. This can however not replace the need for real experiments.

As far as I can see, physicists have not abandoned the search for empirical tests, when working on their theories of quantum gravity. Hopefully, there will be a real breakthrough in quantum gravity research. In the meantime, scientists pursuing research in quantum gravity, should express a large degree of humility and not claim success prematurely. They need to be honest about the, at present quite speculative, nature of their endeavors. If they would do otherwise and no longer pursue empirical results and present their results as if they were already part of established science. In that case I would not find it unfair to view their work as pseudo-scientific.
References


R. Frigg & I. Votsis (2011). ‘Everything you always wanted to know about structural realism but were afraid to ask’. *European Journal for Philosophy of Science* 1:227–276.
244:77–202.


K. Matsubara (work in progress, a). ‘The many worlds of physics and philosophy’.

K. Matsubara (work in progress, b). ‘Thought experiments and quantum gravity’.


Underdetermination Should We Take Seriously’. Philosophy of Science:
Proceedings 68:1–12.

V. J. Stenger (2011). The Fallacy of Fine-Tuning: Why the Universe is Not Designed
for Us. Prometheus Books.

In F. Suppe (ed.), The Structure of Scientific Theories. University of Illinois Press,
2 edn.

D1:1182.


Universe or Multiverse. Cambridge University Press. Originally posted on arXiv
in 2003.

M. Tegmark (2003). ‘Parallel Universes’. In J. D. Barrow, P. C. W. Davies, & C. L.
Harper (eds.), Science and Ultimate Reality: Quantum Theory, Cosmology and
Complexity. Cambridge University Press.

M. Tegmark (2007). ‘The multiverse hierarchy’. In B. Carr (ed.), Universe or
Multiverse. Cambridge University Press.

38:101–150.

University Press.

University Press.

7:667–698.


G. Veneziano (1968). ‘Construction of a crossing symmetric Regge-behaved

D. Wallace (2008). ‘Philosophy of quantum mechanics’. In D. Rickles (ed.), The
Ashgate Companion to Contemporary Philosophy of Physics. Ashgate Publishing
Limited.

Review Letters 59:2067.


and Quantum Gravity 23:4231–4236.

B443:85.


E. Witten (2004). ‘Perturbative Gauge Theory As A String Theory In Twistor Space’.
Communications in Mathematical Physics 252:189–258.