PROBLEMS IN APPLYING MATHEMATICS: ON THE INFERENTIAL AND REPRESENTATIONAL LIMITS OF MATHEMATICS IN PHYSICS.

by

Kevin J. Davey

B. Sc., Monash University, 1992

M. Sc., University of California at Los Angeles, 1996

M. A., University of Pittsburgh, 1999

M. Sc., University of Pittsburgh, 2003

Submitted to the Graduate Faculty of

the Faculty of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2003

UNIVERSITY OF PITTSBURGH

FACULTY OF ARTS AND SCIENCES

This dissertation was presented

by

Kevin J. Davey

It was defended on

October 24, 2003

and approved by

John Norton, History and Philosophy of Science

John Earman, History and Philosophy of Science

Laura Reutsche, Philosophy

Ken Manders, Philosophy

Mark Wilson, Philosophy

Robert Geroch, Physics (University of Chicago)

Dissertation Director: John Norton, History and Philosophy of Science

PROBLEMS IN APPLYING MATHEMATICS:
ON THE INFERENTIAL AND REPRESENTATIONAL LIMITS
OF MATHEMATICS IN PHYSICS.

Kevin J. Davey, PhD

University of Pittsburgh, 2003

It is often supposed that we can use mathematics to capture the time evolution of any physical system. By this, I mean that we can capture the basic truths about the time evolution of a physical system with a set of mathematical assertions, which can then be used as premises in arbitrary mathematical arguments to deduce more complex properties of the system.

I would like to argue that this picture of the role of mathematics in physics is incorrect. Specifically, I shall assert :

The Deduction Failure Thesis: Bodies of knowledge in physics are generally not closed under otherwise valid mathematical argument forms.

The Representation Failure Thesis: We cannot assume that the state of any system, together with its fundamental laws, can be captured by some set of mathematical assertions or equations. In fact, it is more likely that the world is not representable by a set of mathematical assertions or equations than that it is.

The dissertation largely consists of arguments for these two theses.

TABLE OF CONTENTS

Appendices

1.0 INTRODUCTION AND PROLOGUE TO PART 1

1.1 THE FALLIBILITY OF MATHEMATICS.

A scientific theory presents us with a set of basic laws about the world. Often, these basic laws are rather abstract, and not *directly* useful in explaining and predicting phenomena of interest. With the help of mathematics, however, we may often deduce more complicated facts of greater utility that follow from these basic laws. Indeed, the usefulness of mathematics in science is often taken to lie in precisely this fact – that it can help us identify the more complicated, useful facts that follow from the sometimes abstruse basic laws from which we begin.

For instance, classical electrodynamics consists of Maxwell's equations:

$$\nabla \cdot E = \rho/\epsilon_0$$

$$\nabla \times E + \frac{\partial B}{\partial t} = 0$$

$$\nabla \cdot B = 0$$

$$\nabla \times B = \mu_0 J + \frac{\partial E}{c^2 \partial t}$$

along with the Lorentz force law:

$$F = q(E + v \times B).$$

On their own, these basic assertions are not very useful. If we are interested in finding out what light will do when it strikes a body of water, the laws given above do *not* immediately provide us with an answer. What we must do is put these laws together in complicated ways, using various mathematical techniques, to develop a theory of how electromagnetic waves behave at the interface between dielectric media.

One might think that what makes all this possible is the fact that if we are committed to a some body of scientific facts, then we must be committed to all the mathematical consequences of that body of facts. This claim is so important, that we shall give it a name:

The Deduction Thesis: We are committed to the mathematical consequences of any body of scientific facts to which we are committed.

By a 'body of scientific facts', I mean a set of propositions sufficient for making specific predictions. I talk about 'bodies of scientific facts' rather than 'theories', because I want to include initial conditions, the values of physical constants etc., amongst the body of scientific facts to which a scientist may be committed at any moment. So for instance, in order to get any mileage out of Maxwell's equations and the Lorentz force law, one needs to supplement them with facts like: electrons have a charge of $1.6 \times 10^{-19} C$, this wire carries a linear current of $50A$ in the z direction, and so on. Such facts, when they obtain, are to be included in the 'body of scientific facts' to which a physicist may be committed in a given calculation.

The attraction of this thesis is clear – it provides us with a simple explanation of how we can do so much in science with so little, given the rich argumentative resources of mathematics.

In spite of this, I think the Deduction Thesis is false. Let us focus on the case of physics. I claim that the Deduction Thesis is true of practically *no* body of knowledge in physics – be it concerned with classical mechanics, classical electromagnetism, quantum mechanics, thermodynamics, and so on. I think that practically every successful body of knowledge in physics that has ever been constructed has violated the Deduction Thesis in one way or another. As an *empirical* claim about physics, the Deduction Thesis fails badly. Thus, I advance the following thesis:

The Deduction Failure Thesis: Bodies of knowledge in physics generally violate the Deduction Thesis.

That is to say, physicists do *not* necessarily take themselves to be committed to all the mathematical consequences of bodies of knowledge that they are in turn committed to.

This is not intended to contradict the fact that mathematical deduction plays an indispensable role in physics. But that, on its own, does not entail the Deduction Thesis. It is possible that mathematics is an indispensable tool for extracting usable information from physical facts, *without* it being the case that mathematics is *infallible* in this capacity. The precise role of mathematics in physics is far subtler than indicated by the Deduction Thesis. It will be part of the job of this dissertation to explain how, and in doing so defend Deduction Failure Thesis.

One possible reaction to all this is to suggest that the Deduction Failure Thesis is simply a symptom of (i) our ignorance of the final theory of physics, and (ii) our ignorance of the detailed, microscopic states in which physical systems find themselves. Given such ignorance, we should not be surprised that our desperate and clumsy attempts to theorize about the world should result in theories with awkward and ugly features, such as a failure to be closed under otherwise valid mathematical inferences. Once such ignorance is eliminated, however, we should be able to theorize and philosophize about the universe without worrying that an otherwise valid mathematical inference might take us from true premises to a false conclusion.

Implicit in this response is the idea that there is *some* set of mathematical equations and assertions that describes reality exactly – even though, of course, we do not know what those equations and assertions are. This idea is also implicit in much philosophical work on the nature of physical reality – see, for instance, Russell's famous essay on causation [81]. The philosopher who does not want to concern himself with the gaps in man's knowledge of the physical world therefore often blindly assumes that *some* set of mathematical assertions and equations describes reality exactly, and presses on from there. Such a view seems to have been taken for granted by the Pythagoreans, Kepler, Galileo, Gauss and Einstein (as we shall see in Chapter 6.)

However, I think that this view of the relationship between physics and mathematics – according to which the Deduction Failure Thesis is simply a symptom of our relative ignorance about the world, which can be overlooked for certain purposes by assuming that *some* set of mathematical equations and assertions captures reality perfectly – is nothing other than an unjustified expression of optimism about the expressive power of mathematics.

Such a view takes for granted that the fundamental workings of the universe can be captured by mathematical assertions or equations, and that we can begin from such assertions (if we knew them) to draw conclusions about the world in just the same way that a mathematician begins with a set of axioms and derives more complicated theorems. But I do not think that such a world-view can be taken for granted. In fact, I think that (in several senses to be defined later), it is much *less* likely that the fundamental workings of the universe can be captured in purely mathematical terms than not. Thus, I also advance the following thesis:

The Representation Failure Thesis: We cannot take for granted that the world is representable in mathematical terms – that is to say, we cannot take for granted that the state of the world, together with all its fundamental laws, can be captured perfectly by some set of mathematical assertions or equations. In fact, it is more likely that the world is *not* representable in mathematical terms (in this sense) than that it is.

The Deduction Failure Thesis and the Representation Failure Thesis say two very different things. The former is concerned with the sorts of theories that physicists actually produce, and asserts that validity under arbitrary mathematical inferences cannot be assumed. The later is concerned with the possibility of a type of theory which no actual physicist has ever produced, or may ever produce, and asserts that we cannot take for granted that such a theory even exists – i.e., that we cannot blindly assume that mathematics is capable of perfectly capturing the workings of reality. However, although these two theses say quite different things, they can be used together to paint a novel picture of the relationship between mathematics and physics. Let us see how.

What sort of picture is painted by the conjunction of the Deduction Failure Thesis and the Representation Failure Thesis? Let us consider an analogy. A primitive meteorologist stares at the clouds. He finds it natural (for whatever reason) to describe the properties of clouds anthropomorphically. He describes a certain shape of cloud as 'angry', another as 'peaceful', some as 'lazy' and others as 'energetic'. He then uses generalizations about people to infer facts about clouds – e.g., that angry clouds tend to stay angry unless they come in contact with a peaceful clouds, and so on. Let us summarize all this by saying that he uses an *anthropomorphic language* to theorize about clouds.

Now it turns out, of course, that our primitive meteorologist's anthropomorphic language is not really perfectly suitable for a discussion of clouds. In any explanation of the way in which clouds evolve over time, certain facts about the behavior of people will be helpful, while other facts about the behavior of people must conveniently be ignored. Moreover, the chances of constructing a complete theory of cloud evolution (i.e., a theory which could definitively be said to be the 'final word' on the subject) using only an anthropomorphic language seems very slim. I would summarize this situation by saying that there is only a 'limited congruity' between the meteorologist's anthropomorphic language and his desire to theorize about the time evolution of clouds. This is not to say that the primitive meteorologist's anthropomorphic language is *unsuitable* for an analysis of clouds – to the contrary, it might be so suited that the thought of abandoning the anthropomorphic language would seem like folly to any practical-minded person. But regardless of this, the limitations just described on the ability of an anthropomorphic language to perfectly capture the evolution of clouds, suggests that there is only a 'limited congruity' between the two.

I would like to make the same sort of claim about the use of mathematics as the language for physics. Mathematics is an extraordinarily helpful language for describing reality. But if the Deduction Failure Thesis is right, valid mathematical arguments must sometimes conveniently be ignored when theorizing in physics. Likewise, if the Representation Failure Thesis is right, we cannot take for granted that the workings of the universe can be expressed in purely mathematical terms. Thus, there is only a 'limited congruity' between the physicist's mathematical vocabulary, and his desire to perfectly describe physical reality. Of course, this is not do deny that, for practical purposes, the language of mathematics is *excellent* for discussions of physical reality – so much so that it is hard to imagine any practical minded person doing good physics in anything other than mathematical terms. But in spite of this, if the Deduction Failure Thesis and the Representation Failure Thesis are correct, there are limitations on the capacity of mathematics to fully capture physical reality. This is the picture of the relationship between mathematics and physics that emerges from the Deduction Failure Thesis and the Representation Failure Thesis.

Part I of the dissertation will be concerned with the Deduction Failure Thesis, while Part II will be concerned with the Representation Failure Thesis. Because this introductory chapter also serves as a prologue to Part I, much of the rest of this chapter will revolve around a discussion of the Deduction Failure Thesis.

1.2 THE MEGARANS.

Doubts about the Deduction Thesis are not entirely new in the history of philosophy. It will be interesting to briefly consider one particular historical precedent, before moving on.

The Deduction Thesis itself goes back to Plato and Aristotle – for instance, it is implicit in many of the logical works of Aristotle, such as the Posterior Analytics. The philosophers of Megara,[1] however, such as Euclidies, Eubulides and Stilpo, were opposed to many of the beliefs of Plato and Aristotle. Their main arguments against Athenian philosophy consisted of succinct paradoxes, which were supposed to show the futility of the sort of reasoning found in Aristotelian science. For instance, the paradox of the heap ('Does one grain of wheat make a heap? Do two grains? Add a grain and there is yet no heap; when does a heap begin?') and the Liar Paradox ('If a man acknowledges that he lies, does he lie or speak the truth?') originated from the Megarans, and were intended to show difficulties with the Aristotelean style of argument. Because the Megarans' arguments consisted of short, pithy contradictions such as these, their school was subsequently called the 'Eristics', from the Greek 'eristikos', meaning to strive or dispute.

It is difficult to know precisely what broader point the Megarans were making with their paradoxes, in part because none of the works of the Megarans survive. Most of our knowledge of the Megarans comes from second hand sources, such as Aristotle. History being a story that is told by the victors, the Megarans are therefore often presented as verbal tricksters, victims of confusion, having no real systematic point to make against the Aristotelean system, and given very little attention in most presentations of ancient philosophy.

However, I would like to suggest that the sorts of paradoxes considered by the Megarans may well have been thought of as evidence against something like the Deduction Thesis, as advanced by Plato and Aristotle. The Megarans thought (rightly or wrongly) that even impeccable reasoning was capable of leading man astray. (This is not to say that the Megarans

[1] Megara was a city neighbouring Athens.

thought that knowledge was impossible, or worthless – to the contrary, they agreed with the Socratic claim that virtue was knowledge, and that it was knowledge that distinguished the wise man from the fool.) The Megaran philosophy therefore seemed to revolve around the idea that while reasoning was the source of all knowledge (and even the source of virtue), it was not *infallible*, even when one began from the truth. This seems similar to what it might mean to deny the Deduction Thesis.

In this dissertation, I will not be trying to defend the Megaran philosophy – I will not mention the paradox of the heap or the liar paradox again. I mention the Megarans because there is a certain similarity in what they set out to prove and what I am setting out to prove. There are, however, also significant differences: my concern is explicitly with mathematics, while the Megarans were concerned with a more general critique of the use of reason in science. In addition, the type of problems that arbitrary mathematical inferences can lead us to are quite different from the 'paradoxes' with which the Megarans were concerned. Nevertheless, the Megarans serve as an interesting comparison for the present project.

1.3 THE DIFFERENCES BETWEEN MATHEMATICIANS AND PHYSICISTS.

Let us return to modern times, and to the Deduction and Deduction Failure Theses presented earlier.

If the Deduction Failure Thesis is correct, physicists more or less systematically commit themselves to bodies of knowledge that violate the Deduction Thesis. Assuming this is correct, it is natural to ask *why* physicists consistently come up with bodies of knowledge that are not closed under arbitrary mathematical inferences. Is this an accident, or a symptom of some underlying philosophical attitude that physicists have about the role of mathematics in physics?

I would like to argue that the latter is the case – that physicists have a certain well-formed and defensible attitude about the role of mathematics in physics, that guides them in the discovery process, and has a tendency to cause them to commit themselves to bodies of knowledge that are not closed under arbitrary mathematical inferences. In subsequent chapters I will investigate this attitude in detail, but it will be helpful to make some preliminary remarks now.

I want to try and understand what this attitude could be that leads physicists to commit to bodies of knowledge that violate the Deduction Thesis. In order to do so, I shall first say a few things about the physicists' attitude to mathematics quite generally.

At first glance, there appears to be great unity between the projects of mathematics and physics. But beneath this veneer of unity lie deep cultural differences. Consider the following three facts:

(I.) What the physicist takes to be a persuasive argument will often leave the mathematician quite unpersuaded. The physicist will freely introduce approximations, interchange integral signs, derivatives, and do whatever is needed in order to get his calculations to work. Sometimes, he will use poorly defined, or even inconsistent mathematical entities (such as path integrals or delta functions) at crucial points in his arguments. (See [90] for some interesting examples of this.) The mathematician, on the other hand, works almost exclusively with precisely defined objects and an extremely high standard of argumentative rigor, and thus finds the arguments of the physicists difficult to penetrate at best, and utterly unconvincing at worst.

(II.) Arguments that the *mathematician* finds persuasive will often leave the *physicist* unconvinced. The physicist generally shies away from mathematical arguments with dubious connections to the concrete workings of nature, often viewing such results with suspicion. For instance, some physicists are skeptical about whether the literature on stable, periodic solutions to the N-body problem (see [62], for example) really tells us anything about the physical world. Other physicists are skeptical about whether certain approaches to the problem of the radiation reaction in classical electrodynamics (see [51], for example) take the notion of a 'point-charge' too seriously. These examples demonstrate that physicists can sometimes reject the putative physical consequences of mathematically valid arguments, out of a concern that the mathematics is somehow leading them astray, or that some sort of idealization is being taken too seriously, or for any one of a host of other reasons.

(III.) When physicists and mathematicians *do* agree that some mathematical argument appears to have real physical significance, the physicist will sometimes announce that the effectiveness of mathematics is 'unexpected', 'surprising', or even 'unreasonable'. (Most notably, see Wigner's [101].) That formal methods should be effective in mathematics is, of course, unsurprising — yet to some physicists, the effectiveness of formal methods in *their* discipline *is* regarded as a surprise. Wigner makes this point dramatically when he says *'... the miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve.'*

Of course, not all physicists view mathematics in the way just outlined in these three points. Much diversity of thought exists within the physics community on the issue of the precise relationship between mathematics and physics. There are physicists who hold some, but not all, of the views I have just described, and there are physicists who hold none of them.

Nevertheless, what I have sketched above is a *general* attitude about mathematics that is very much alive in the physics community today. The existence of this attitude demonstrates that the relationship between mathematics and physics is not as tight as the outsider might naively think, and reveals a deep culture difference between mathematicians and many physicists. (We will also eventually see that this attitude is closely connected with the Deduction Failure Thesis.) What are we to make of all this?

1.4 THINKING ABOUT THE CULTURE DIFFERENCE.

One response to the cultural difference outlined above is to suggest that the physicists deserve a great deal of blame for the situation. That the arguments of the physicists are not persuasive by the mathematicians' standards, for instance, could simply be viewed as a failing on the physicists' part. Likewise, the physicists' suspicion at 'unphysical' arguments is equally unfounded – valid mathematical arguments are valid mathematical arguments, regardless of the shape those arguments take. The burden is surely on the physicist to explain why such results are physically irrelevant. Finally, it might be argued that a sense of surprise at the effectiveness of mathematics in physics must rest on a faulty philosophy of mathematics that the physicist has somehow acquired. Behaviors **(I.)**–**(III.)** are thus deficiencies on the physicists' part. The view that physicists are largely to blame for the culture difference I have been outlining is perhaps best summarized in Hilbert's famous proclamation (see [104]) that *'physics has become far too difficult to leave to the physicists.'*

The attitude just described is quite harsh. One might imagine a more moderate view, according to which the physicist is given a little more latitude when exploring developing areas of physics. For instance, perhaps it is acceptable for a physicist to traffic in unrigorous arguments, or to be unpersuaded by arguments that are completely rigorous, when the area of physics in question has not yet reached firm footing. Likewise, perhaps the physicist is justified in being skeptical (or at least, cautious) about the offerings of mathematics with respect to a new area of physics, out of concern that the mathematicians' offerings may be infected by all sorts of misleading prejudices. According to this view, the behaviors of the physicist described in points **(I.)-(III.)** are perfectly reasonable when dealing with a *fledgling* area of physics – but once our understanding of an area of physics reaches a certain point of maturity, such excuses no longer apply. The harsher, Hilbert-style view described above then becomes appropriate.

But I think there is something unfair even about this more moderate position. Why should the standards of argument that work well in mathematics be forced upon the physicist, even when he is discussing phenomena that are well understood physically? Someone like Hilbert might reply that the standards of clarity and rigor that one finds in mathematics are simply the standards of clarity and rigor that are appropriate to *any* quantitative discussion of reality.

I do not think, however, that such a view is as strong as it might at first seem — even when applied to well understood areas of physics. In fact, I would like to suggest that there is *nothing wrong at all* with the practices and attitudes of the physicist that we find in the three examples **(I.)-(III.)** given earlier. Specifically, I would like to argue that physicists are not necessarily misguided in embracing non-rigorous mathematical methods, and ignoring certain rigorous arguments, even in mature theories. I also claim that the physicist is right to worry about the limits of mathematical expression, and even feel a certain surprise when the tools of mathematics turn out to be fruitful to him. I take the truth of these claims to be independent of whether it is an immature or mature area of physics that is under consideration. The arguments for these claims will fill much of the dissertation.

1.5 THE PRINCIPLE OF INFERENTIAL RESTRICTIVENESS.

Thus far, I have asserted that physicists have a tendency to commit themselves to bodies of knowledge that violate the Deduction Thesis, and I have asked why this is the case. I have suggested that physicists produce such theories because of an underlying view about mathematics that authorizes – and even encourages – them to produce such theories. In order to understand what this view about mathematics might be, I have described a few peculiar features of the physicists' general attitude towards mathematics.

We must now try and move from a discussion of these peculiarities to a more systematic understanding of the attitude of the physicist towards mathematics. What sort of connections exist between behaviours **(I.)–(III.)**? Do they proceed from utterly distinct sorts of concerns, or is there some more general type of concern that motivates them all?

I would like to suggest that there is a philosophical principle about the relationship between mathematics and physics that (i) many physicists would subscribe to (implicitly or explicitly) and that (ii) motivates all three of the physicists' behaviors outlined above. Such a principle makes intelligible and lends unity to the peculiar attitude that many physicists have towards mathematics.

The principle is a principle about the way in which mathematics may be *used* in physics. The principle states that, in physics, mathematics is generally to be used in an *'inferentially restrictive'*, rather than an *'inferentially permissive'* way. I shall now explain what is meant by the distinction between inferentially permissive and inferentially restrictive uses of mathematics.

When the mathematician tries to prove a theorem or solve a mathematical problem, he is free to invoke any concept, technique or theorem from any area of mathematics that he wishes. No *legal* inference from true premises is barred – although to say an inference is not barred is obviously not to say that it bears fruit. It is precisely because of this 'permissive' nature of inference in mathematics that one comes across surprises such as the set-theoretic proof of the number-theoretic claim that transcendental numbers exist. By describing mathematics as *inferentially permissive*, I shall mean that there is no restriction on the concepts or legal rules of inference that one can invoke in attempting to prove a theorem or solve a puzzle in mathematics. In a mathematical proof, any legal inference is permitted, and any concept can be invoked, regardless of the subject matter introduced.

By contrast, I claim that a close look at the methodology of physics shows that the physicist is *not* willing to consider absolutely *any* sort of mathematically legal technique or rule of inference when solving a problem in physics. His willingness to exploit the inferentially permissive character of mathematics only goes so far. In certain cases, a physicist is perfectly happy to reject a mathematical consequence of premises he accepts, claiming that the chain of reasoning in question is 'misleading', 'unphysical', or just plain *false*. Because of this, I shall call the physicist's use of mathematics *inferentially restrictive* – meaning that the physicist does not always feel forced to accept propositions that follow mathematically from propositions he accepts.

Having said this, I can now state the principle that gives unity to the peculiar attitudes that physicists have towards mathematics:

The Principle of Inferential Restrictiveness: In physics, the use of mathematics is generally *inferentially restrictive* – that is to say, the physicist does not feel obligated to endorse the mathematical consequences of facts that he endorses.

Obviously, the Principle of Inferential Restrictiveness and the Deduction Failure Thesis are closely related. The Principle of Inferential Restrictiveness describes the behavior of physicists – specifically, it describes the way in which they treat mathematics. The Deduction Failure Thesis, on the other hand, describes the *result* of this behavior – a tendency to commit to bodies of knowledge that violate the Deduction Thesis. Thus, the Principle of Inferential Restrictiveness explains why it is that physicists systematically commit to bodies of knowledge that violate the Deduction Thesis.

The Principle of Inferential Restrictiveness also explains the three behaviors outlined earlier, as we shall see in subsequent chapters. In particular, it explains why the physicist is sometimes willing to ignore the physical consequences of otherwise mathematically valid arguments (see Chapter 4). It also explains how the physicist is able to sacrifice rigor, without falling into nonsense (see Chapter 3). Furthermore, it explains why there is something surprising when it turns out that we *can* trust reasoning that we had no a-priori right to trust (see Chapter 4, also related are Chapters 7 and 8.) Thus, the Principle of Inferential Restrictiveness lends unity to the curious behaviors cited earlier, and answers our question of why physicists systematically commit to bodies of knowledge that violate the Deduction Thesis.

However, the Principle of Inferential Restrictiveness also raises many questions. If mathematical validity is not enough to ensure that the truth of a set of premises really entails the truth of a conclusion, then what is? If *some* mathematical arguments can't be trusted, then why are *any* mathematical arguments trustworthy? How do we decide which sorts of mathematically valid arguments are trustworthy, and which aren't, and on what basis?

In addition, one might wonder whether, in citing the Principle of Inferential Restrictiveness, we have really provided a satisfying explanation of *anything.* Specifically, one might wonder *why* the physicist is willing to adopt an inferentially restrictive mathematical methodology in the first place. Until this issue is addressed, it might be felt that we have provided only a very superficial explanation of why many physicists treat mathematics in the way that they do, and why they produce theories that violate the Deduction Thesis. In order to tackle this issue, as well as some of the issues discussed in the previous paragraph, we will have to think a bit more seriously about what inferential restrictiveness is, and what makes it attractive.

1.6 INFERENTIAL RESTRICTIVENESS AND SCIENTIFIC DIFFICULTIES.

Why might a physicist feel that not all the mathematical consequences of an otherwise successful body of knowledge are to be trusted? We will consider this question in detail in Chapters 2–4, but I shall make some preliminary remarks now.

Let us consider a couple of examples of difficulties in physics.

(1.) We might naively think of the atoms in a dilute gas as tiny billiard balls which collide elastically. Once we do this, we get to apply the whole machinery (both the vocabulary and rules) of macroscopic, elastic collisions to the dilute gas. This is extraordinarily profitable, and leads, for instance, to the ideal gas law, which is a very accurate empirical result.

However, problems arise at low temperatures – for instance, (i) the theoretically predicted value for the specific heat turns out to be quite different from that observed experimentally, (ii) a gas of electrons at zero degrees Kelvin still exerts pressure on the walls of its container (i.e., $T=0$ and $P\neq 0$), and so on. The solution, of course, is to completely abandon the language of 'nearly elastic collisions' acquired from macroscopic experience, and replace it with the language of quantum statistical mechanics.

(2.) Classical electrodynamics is often developed in terms of charged point particles, and subsequently generalized to solids (capacitor plates, antennae, and so on), with extraordinary success. However, fundamental problems arise when we try to understand the radiation reaction on an accelerating point charge. Depending on the route one takes, one gets infinite forces, effects preceding causes, or similar inconveniences. The consensus view amongst physicists here is that one can only resolve the problem by abandoning the classical language of 'charged point particles', and moving to (relativistic) quantum field theory – specifically, quantum electrodynamics.

In each case, we are presented with body of knowledge, and a difficulty. In both cases, a possible response to the problem is to throw away our old theory, and try to come up with a new theory that explains the anomalous phenomena. When we can see how to do this, this is presumably the route of choice. It is, however, also the most challenging route to take.

What are we to do if we do not yet have the insight to pursue this option? We can continue trying to tinker with the conceptual apparatus we already have in order to get things to work. But in the examples just given, this does not yield fruit – there is no real way to explain, for instance, the deviation in observed values of the specific heat of gases at low temperatures as long as one thinks of a gas as a collection of classical objects elastically colliding; and likewise, there is no real way to satisfactorily calculate the radiation reaction on a charged, accelerating point particle armed only with Maxwell's equations (in classical form). What an honest scientist is (eventually) forced to do is to confess that, in these situations, his body of knowledge runs up against the limits of its usefulness. What is he then to do? He will go on using the same arguments he previously used; but he will be careful not to endorse the conclusions of arguments which take him into areas in which he now knows problems exist. So, for instance, the scientist will continue to use the macroscopic elastic collision metaphor when thinking about gases, but will simply abstain from drawing conclusions about gases at low temperatures. Likewise, he will continue to endorse Maxwell's equations, but will suspend his judgment about arguments involving the radiation and motion of charged point particles. In both cases, he ends up placing some *restriction* on the set of inferences he would ordinarily use to describe the phenomena in question.

In each case then, the physicist ends up with a body of knowledge which is such that as long as he avoids certain troublesome (but otherwise valid) argument forms, true premises will be taken to true conclusions. The physicist continues to say that dilute gases are like collections of billiard balls that undergo elastic collisions, and confidently continues to endorse all the conclusions that follow – except in the case of low temperatures, where he simply shrugs his shoulders, and says that his concepts fail him.[2] Likewise, he goes on talking about Maxwell's equations and the Lorentz force law as he did before – except in the case of accelerating, radiating point charges, where again he simply shrugs his shoulders, and says that the point charge formalism shouldn't be taken *that* seriously.

In both of these cases, the physicist has some sort of 'inferentially restrictive' methodology. For instance, the physicist will *not* necessarily endorse the conclusion of *any* old logically valid argument that proceeds from the macroscopic elastic collision model of gases – even though he will unproblematically accept such arguments *in the right circumstances*. Likewise, the physicist will not necessarily endorse the conclusion of *any* old logically valid argument that proceeds from Maxwell's equations and the Lorentz force law – even though again, he will unproblematically accept such arguments in the right circumstances.

In the examples given, inferential restrictions were adopted upon the discovery of some *actual* anomaly. However, it is also possible to apply the General Methodological Principle *pre-emptively*. For instance, an especially astute physicist might have worried about the extension of the macroscopic elastic collision model of gases to low temperatures long before problems were discovered. Likewise, modern physicists might worry about mathematical formalisms that presuppose the infinite divisibility of space or time, or that require sharp eigenvalues for operators, and so on. If one is inclined to worry about such aspects of our

[2]One might think that, in this case, one can simply restrict the relevant law to certain 'regimes', so that no restriction of inferences is necessary. We shall see in Chapters 2 and 4, however, that this is not so easy.

formalism, then it would be rational to adopt inferential restrictions pre-emptively, even before specific problems are discovered, in order to block arguments that critically exploit the aspects of our formalism about which we might be suspicious. We will discuss this a little more in Chapter 4.

What lessons can we derive from all this? Inferential restrictions enable the physicist to continue using a body of knowledge in cases in which it is beneficial to do so, while allowing the physicist to avoid the dangerous territory into which he knows his theory can force him, if taken too seriously. The adoption of inferential restrictions is a useful (and indeed, rational) thing to do when one is uncertain about the appropriateness of one's conceptual tools. The uncertainty can be caused by an *actual* anomaly, or it can be caused by a hunch that anomalies might be nearby. Either way, it can be perfectly reasonable for a physicist in a certain situation not to trust all the mathematical consequences of an otherwise successful theory.

With the help of some of this, we can give provisional answers to some of the questions posed in the previous section. (These answers will all be developed in further detail in subsequent chapters, as indicated.)

Question: In physics, if *some* mathematical arguments can't be trusted, then how can *any* mathematical arguments be trusted?

Answer: In both of the examples just given, we have a body of knowledge which is not entirely suitable for discussing some broad class of physical phenomena. Nevertheless, the body of knowledge *is* very useful for *certain* discussions of such phenomena. Inferential restrictions are implemented in order that it be possible to (i) take advantage of the strengths of the theory, while at the same time (ii) avoiding the problems into which one can fall by relying on the theory too heavily. So one can trust such mathematical arguments when all inferential restrictions are obeyed, even though one cannot trust *arbitrary* mathematical arguments. Further details on this issue are developed in Chapter 2.

Question: If mathematical validity is not enough to ensure that the truth of a set of premises really entails the truth of a conclusion, then what is?

Answer: When a theory fails in the way seen in the examples given, it means that only certain types of mathematical arguments using that theory are reliable. Which types of arguments these are will depend on the idiosyncrasies of the world. Mathematical validity is not enough to ensure that the truth of a set of premises entails the truth of a conclusion – in addition to mathematical validity, one needs the mathematical argument in question to be one of those to which the world just happens to be suited. Again, further details on this are developed in Chapter 2, where I suggest that we think of Poset Homomorphisms as 'facts about the world' which the scientist must discover.

Question: How do we decide which sorts of mathematically valid arguments are trustworthy, and which aren't, and on what basis? Can we ever know for sure that a given form of mathematical inference is completely trustworthy?

Answer: We will explore some particular cases of inferential restrictiveness in the subsequent chapters; we will see that the decision of which inferential restrictions we adopt are largely based on experience and conjecture, rather than any straightforward general principle. Also, much like any other assertion in science, it is unlikely that assertions about which inferential restrictions are appropriate can be known with certainty – even though this is not to deny that there are facts of the matter about such things. Again, we will elaborate on this in Chapter 2, as well as subsequent chapters.

1.7 OUTLINE OF PART 1.

Having introduced the main Theses (the Deduction Failure Thesis and the Representation Failure Thesis) that I will be defending in this dissertation, and having sketched some of the ideas that are to be developed in subsequent chapters, I shall make a few comments about the structure of the dissertation itself.

Part 1 (Chapters 2–4) of the dissertation will consist of a defense of the Deduction Failure Thesis, while Part 2 (Chapters 5–8) of the dissertation will consist of a defense of the Representation Failure Thesis. I shall outline the contents of Part 1 here; saving the outline and comments on Part 2 for later.

1.7.1 Chapter 2: Approximations and Idealizations.

The practical calculations of physics generally involve idealizations of some sort or other. In typical calculations, we assume that wires are negligibly thin, that stars are spheres, that electrons are point charges, and so on. When we try to correct for these oversimplifications, all we generally do is replace one idealization with another more sophisticated idealization. Knowing how to reason with idealizations is therefore an indispensable part of the physicist's repertoire.

Reasoning with idealizations involves accepting as *true* the conclusions of arguments whose premises may be, as they stand, literally *false*. For instance, a physicist might explain the order of magnitude of the attraction experienced by two parallel current-carrying wires by referring to the standard application of Ampere's law to a line of current, even though the physicist knows full well that the wires before him are not literally '*lines* of current' at all. What this means is that physicists are often persuaded by arguments whose premises they recognize to be false. One might think that there are simple ways to 'sanitize' arguments like the one involving Ampere's law, so that they consist only of true premises. But I argue that such attempts do nothing but replace one idealization with a more sophisticated idealization, so that at the end of the day, we are still left with a (persuasive) argument with false premises.

I think that the only way to understand how physicists reason in such cases is to view them as adopting an inferentially restrictive methodology. When a physicist makes an idealized assertion about the world, his trust in it is limited – he will use it as a premise in some sorts of arguments, but in other sorts of arguments he will not be so willing to use it. Thus, the (necessary) use of idealizations in physics provides us with a piece of evidence for the Principle of Inferential Restrictiveness, according to which physicists do not feel obligated to endorse the mathematical consequences of facts that they endorse.

I also argue that in such cases, there is a type of *justification* physicists can give for endorsing the conclusions of some arguments, while rejecting the conclusion of others, given some fixed set of endorsed premises. (This is the Poset Homomorphism Result.)

In addition, I shall argue that in many cases in which idealizations are involved, physicists are expressly *prohibited* from endorsing the conclusions of otherwise mathematically acceptable arguments, all of whose premises they endorse. They have no choice but to reason inferentially restrictively. This gives us an important piece of evidence for the Deduction Failure Thesis, according to which the bodies of knowledge to which physicists commit themselves are generally *not* closed under arbitrary mathematical consequence.

1.7.2 Chapter 3: Mathematical Rigor.

The next phenomenon I wish to consider is that of non-rigorous mathematical reasoning in physics. Physics textbooks often contain arguments that do not pass muster with the mathematicians. Sometimes, this is nothing serious – details can be filled in by the reader who has a fetish for the highest standards of mathematical rigor. But things are not always so easy. In some cases, it is *not* obvious how to fill in the details, and in other cases, it is clearly *impossible* to fill in the details, insofar as inconsistent demands are being put on some putative mathematical object. What are we to make of such situations?

This is a hard question to attack in full generality, because there are so many different ways in which the mathematician's standards of rigor can be violated by the physicist. For the purposes of my argument, it suffices to fix one particular example of non-rigorous mathematics that one finds in physics. I shall focus on the case in which one uses a mathematically ill-defined, or incoherent concept in order to derive some specific physical result. For instance, no function satisfies the properties of the delta function, and no measure satisfies the various requirements one might impose to try and make sense of the Feynman path integral (in a measure-theoretic sense). Nevertheless, delta functions and path integrals occupy a fundamental role in the day to day calculations of many physicists.

How are we to understand what physicists are doing when they reason unrigorously in this way? My suggestion is that we should view physicists in such moments as adopting an inferentially restrictive mathematical methodology – that is, as not feeling *bound* to accept the conclusion of any valid mathematical argument, all of whose premises they accept. By adopting such inferential restrictions, the physicist gets to enjoy and exploit the properties of a much wider class of mathematical object than he would if he outright refused to admit incoherent mathematical objects into his vocabulary. Thus, we shall see that the use of non-rigorous mathematical reasoning in physics provides us with another piece of evidence for both the Principle of Inferential Restrictiveness and the Deduction Failure Thesis.

I also show how justificatory questions about non-rigorous arguments can be addressed with the help of Poset Homomorphism Results, in analogy with Chapter 2.

1.7.3 Chapter 4: Unphysical Arguments.

Finally, I wish to consider the fact that physicists are sometimes willing to dismiss the conclusions of perfectly coherent bodies of mathematical reasoning as 'unphysical'. For instance, much work concerning stable, periodic solutions to the N-body problem in classical mechanics is generally viewed by many physicists as having dubious physical content. Likewise, the treatment of singularities and black holes in general relativity were originally viewed by some as unphysical (even though the modern attitude is, of course, different.) Of philosophical interest is the question of whether a physicist can ever be *justified* in having such suspicions about an otherwise mathematically valid argument. (Of course, judgments about whether a body of mathematical arguments are 'physical' will not be *infallible*, as the case of black holes shows. But in spite of this, one might wonder whether it can ever be reasonable for a physicist to at least *conjecture* that such a body of arguments ought to be dismissed as unphysical.)

I shall argue that the best way to make sense of such suspicions is by viewing them as an expression of the fact that physicists do not necessarily feel obligated to endorse the mathematical consequences of other facts that they endorse. (This is the Principle of Inferential Restrictiveness.) In particular, I shall argue that when a physicist expresses reservations about the 'physicality' of some novel type of mathematical argument, he is questioning whether such arguments violate known or unknown inferential restrictions associated with some body of knowledge. I will show that such concerns are rational, by rooting them in facts about Poset Homomorphism results, in analogy with the previous chapters.

In fact, thinking about such phenomena help us to paint a richer picture of the role of mathematics in scientific discovery. I shall argue that when a physicist verifies a law, what he really verifies is the efficacy of a limited class of mathematical techniques, applied to some abstract assertion or law. Whether novel mathematical techniques applied to the same assertion can be trusted therefore remains a genuinely open question. In this way, I shall show that the Principle of Inferential Restrictiveness is a rational principle that is built into the very process of discovery in physics.

The physicists' concerns about 'unphysical' mathematical arguments therefore provides us with additional evidence for the Principle of Inferential Restrictiveness. In cases in which we can actually argue that otherwise acceptable mathematical arguments are unphysical, we also have further evidence for the Deduction Failure Thesis, as I shall argue.

2.0 APPROXIMATIONS AND IDEALIZATIONS.

2.1 INTRODUCTION.

Idealizations – that is, highly simplified representations of reality – play an indispensable role in physics. There are a number of reasons for this. In some cases, calculations are too difficult if simplifying assumptions are not made. In other cases, an analysis is *impossible* if simplifying assumptions are not made. In further cases, the physicist's knowledge of a system is incomplete, and in need of supplementing; naturally, he will try to supplement it in the simplest way possible - without compromising things too much - in order to make subsequent calculations as easy as possible.

The nature of idealizations, and the epistemological problems they raise, have not been ignored by philosophers of science. In [76], Redhead discusses the use of idealizations (which he calls models), and discusses some problems that can arise when they are mishandled. The dependence of physics on idealizations has also been a major theme for Cartwright, especially in [17]. Other philosophers (for instance, Barr [6], Nowak [68] and Weston [99]) have tried to develop more formal accounts of how it is possible to reason with idealizations.

One of the obvious epistemological problems with idealizations is to understand how an argument with idealized – and therefore false – premises can be persuasive. There are various ways of tackling this problem. Roughly speaking, most approaches take an argument with idealized (and therefore false) premises, and 'sanitize' it; replacing the argument with a similar one involving only true premises. For instance, the premise containing the idealization might be replaced by one saying that the physical system is obtainable by a 'small perturbation' from the idealized system. Standard mathematical techniques then kick in. (A very simple example of this approach appears in Weston ([99, 100]).) The net result is an argument with only true premises, yielding the desired conclusion. The epistemological problem of understanding how an argument with false premises can be persuasive is therefore not so much solved, as avoided.

I shall argue, however, that any attempt to deal with idealizations by means of 'sanitizing' is misguided. I shall claim that many arguments in physics involving idealizations *necessarily* have false premises. The presence of false premises in such arguments cannot be eliminated entirely – even though, like a lump under a carpet that can be spread out or pushed into various corners, the false premise in such an argument can often be pushed around in various ways. We must therefore deal with the problem of false premises in a different way.[1]

The situation, however, is not as bad as one might think. I shall show there are simple techniques for reasoning with false premises, and that these techniques save the day when it comes to applying physics to reality. Specifically, I shall argue that it is by adopting inferentially restrictions that the physicist is able to reason successfully with falsehoods – but that without inferential restrictions, he is lost. This will be an important piece of evidence (I shall argue) for the Principle of Inferential Restrictiveness, and the Deduction Failure Thesis.

In section 2, I will motivate my main questions in further detail by focusing on a particular example. I shall present a simple situation in which a physicist uses an idealization to persuade himself of some particular result. I shall pose the question of how such an argument can be persuasive, given that it has false premises, and shall show how attempts at sanitizing the argument do not work. In section 3, I shall then show how inferential restrictiveness saves the day. Section 4 will examine some famous claims of Cartwright in light of the results of the previous sections. In Section 5, I shall conclude by examining the question of how we can view the conclusion of an argument with false premises as 'justified'.

[1]This is not to say that physicists are doing something *wrong* when they use perturbation theory. In the case of perturbation theory, we shall see that the false premise is simply moved elsewhere, in a way that makes subsequent calculations more convenient for the physicist. There is a justificatory question that perturbation theory leaves unanswered, but insofar as this question is largely philosophical, the physicist cannot be criticised for ignoring it.

2.2 CONNECTING THEORY AND OBSERVATION.

2.2.1 The Problem Defined.

Let us begin with an example. Imagine that in the course of trying to understand some piece of equipment, a physicist must calculate the force that two straight wires with known currents running through them exert on one another. If the physicist uses the standard textbook methods, he will treat the charged wires as infinitely thin line segments of uniform current density, and derive the magnetic field strength of each wire using Ampere's law in the usual way. Let us imagine that, by doing this, our physicist obtains a theoretical result agreeing well enough with what he observes in the wires.

Experience suggests that the physicist's calculations will agree with his observations only to a few significant figures. In ordinary circumstances, even two significant figures would be an outstanding achievement. This sort of agreement is more than enough for most practical purposes, and is sufficient for us to declare that the principles of physics have enjoyed yet another success.

Indeed, the physicist should *expect* only approximate agreement with his measured results, as the wires are not truly linear, they are certainly not infinitely thin, the current density is not uniform, and so on. Except for the most controlled of situations, the physicist should probably see little reason to expect agreement to even the 3rd or 4th significant figure of his calculation.

In this case, in order for the physicist to make predictions about the real world, he must use theoretical arguments with premises that he knows to be false. He must assume the wires are infinitely thin, straight, and so on. This is not just a question of convenience – our physicist has highly incomplete information about the system before him, and so has *no choice* but to make some idealized (and therefore false) assumptions about the wires.

This prompts a philosophical question, that will occupy our attention for much of this chapter. By performing some calculations, our physicist was persuaded that his measurement of the attraction between the two wires should have some value, to a decimal place or two. And surely he was right in being so persuaded. But the argument that persuaded him was one with false premises. How can our physicist be right to be persuaded by an argument with premises that are not just false, but *known* to be false? This is the problem that I wish to focus on.[2]

2.2.2 False Starts.

Let us consider a few reactions to the problem just posed.

One might imagine the physicist telling us that his premises are at least 'sort of' true, even if not literally true, and that this is good enough for him. This response, although not necessarily inaccurate, is not philosophically satisfying, because it leaves us wondering what makes a proposition 'sort of' true, and how this strange sort of truth can be employed in persuasive arguments.

Let us consider a different response. It might be suggested that the system in the physicist's lab can be obtained as a 'perturbation' of the idealized system considered in the physicist's theoretical argument, and that because of this, with a little help from mathematics it can be shown that predictions based on the idealized system can be depended upon to a couple of significant figures. Thus, one is able to replace the offending argument by an argument with only *true* premises. Let us call this the 'idealization free' proposal.

There are two problems with the idealization free proposal. I shall consider the first in this subsection, and the second problem in the next subsection.

First, the idealization free proposal simply dodges the issue. It does so by taking an argument with premise:

(P1) Wires W_1 and W_2 are line segments of uniform current density.

and 'sanitizing' it, replacing it by an argument with premise:

(P1′) Wires W_1 and W_2 may be obtained from line segments of uniform current density by a small perturbation.

Thus, when asked how an argument like that beginning with **(P1)** can be persuasive, given that it has premises known to be false, the idealization free proposal simply instructs us

[2]Of course, there are similar cases in which the physicist would *not* be right to be persuaded by such a calculation, if the idealizations involved were so outrageous as to threaten even the first or second significant figures of the theoretical calculation. I assume, however, that the present example is not such a situation.

to consider a different argument beginning with the *true* premise **(P1′)**, and trusts that we will notice how it leads in a more rigorous way to the same conclusion. Our question, however, was not whether the conclusion of the argument beginning with **(P1)** was *true* – our question was rather how it is that an argument beginning with **(P1)**, a premise known to be false, could be the sort of thing which a physicist ought to be persuaded by. Replacing the argument by another only amounts to dodging the question.

In response to this, it might be declared there is something misguided about the philosophical problem I have posed, and that the argument with premise **(P1)** is not *supposed* to be persuasive; instead, it functions as shorthand for an argument beginning with **(P1′)** that presumably *is* persuasive. To demand an explanation for how an argument with premise **(P1)** can be persuasive is like demanding a soufflé from a cookbook — while the book itself can't give you what you want, it can show you how to go about getting it, given sufficient resources and expertise.

But I don't think this refutation of the problem works. First, I think it gets the psychology of persuasion in physics completely wrong. I think that physicists *are* persuaded by arguments beginning with premises like **(P1)**, and that to deny this is to paint an unrealistic picture of persuasion in physics. Specifically, physicists often use perturbation theory to justify arguments that they had *already* been persuaded by, rather than to produce a type of persuasiveness that was previously lacking. This is not to deny, however, that sometimes perturbation theory *can* be used to help us get clear on cases in which we are not sure whether an idealization is harmful. My main point, however, is just that physics textbooks contain many persuasive arguments long before the chapter on perturbation theory.

To this, of course, it could well be replied that physicists should *not* be persuaded by arguments beginning with **(P1)**, but should only be persuaded by arguments beginning with premises such as **(P1′)** – common practice notwithstanding. Such revisionist policies, I think, should give the philosopher of science pause. But regardless of one's views on what it is appropriate or inappropriate for philosophy of science to tell the physicist, this revisionist policy is far more difficult to carry out that it first appears – which brings us to the second main objection to the idealization free proposal.

2.2.3 More false starts.

The most significant problem with the idealization free proposal is that it is *not possible* to write down a rigorous argument with *only true premises* that has the conclusion that there will be a given attraction between the wires (even allowing for some margin of error). One can perhaps imagine such an argument using the perturbation theory of classical electrodynamics – but classical electrodynamics, strictly speaking, is false, and so this will not do if our argument is to have only true premises. One could give an argument using the full machinery of quantum field theory – but again, quantum field theory (in its present form) cannot account for gravitational effects, and so even it cannot give us what we want.

The problem is that if we want to write down an argument with only true premises that leads us to the desired conclusion, we must first know the exact laws of physics. So the strategy of 'santizing' arguments that contain premises like **(P1)** *cannot* work in a state of incomplete knowledge about the ultimate laws of nature. Thus, if one thinks that what makes arguments involving idealizations persuasive is that they can be sanitized (i.e., transformed into arguments involving only true premises), then one is forced to maintain there can be essentially no persuasive arguments in physics until physics is complete. But such a view is obviously absurd. Recognizing the absurdity of such a view, we are forced to admit that it must be possible for persuasive arguments in physics to have false premises. Thus:

Thesis: Persuasive arguments in physics can (and generally do) have false premises.

Can this thesis be avoided? I will consider five ways in which one might try to do so; four in this section, and one in the next.

First, one might try to obtain an argument with only true premises by weakening one's claims about what the laws of nature are. Thus, in our example, instead of deriving the behavior of the wires from Maxwell's equations (and in particular, Ampere's law), one derives the behavior of the wires from the claim that the world largely behaves *approximately as if* Maxwell's laws were true. Thus, one's argument begins with the premises:

(P0′) The world behaves roughly as if Maxwell's laws were true.

(P1′) Wires W_1 and W_2 may be obtained from line segments of uniform current density by a small perturbation.

from which one derives the force of attraction in the straightforward way. Both premises **(P0′)** and **(P1′)** appear to be true. Thus, we have successfully sanitized our original argument, and therefore refuted the thesis given above.

The problem, however, is that **(P0′)** is *not* true in any straightforward sense. Given Maxwell's equations as they stand, bound electrons spiral into nuclei at alarming speeds (on the order of $10^{-8}s$), and all atomic structure becomes unstable – a phenomenon with very observable consequences at all levels. The claim that the world looks roughly as if Maxwell's equations were true is unambiguously *false*. Rather, Maxwell's equations are useful only if one *already* agrees to lie about the world – to pretend, for instance, that atoms have no internal structure. Thus, **(P0′)** is true only if, by 'the world', one means *not* the actual physical world, but rather some idealized description of reality – specifically, if one means by 'the world' something like: 'the way the world would be if matter had no internal structure (amongst other things).' Thus, the argument with premises **(P0′)** and **(P1′)** cannot be viewed as a successful attempt at sanitizing – the falsehood involved in this argument has simply been pushed to another place under the carpet.

Let us consider a second attempt. We replace **(P0′)** above with **(P0″)**:

(P0″) Certain objects (for instance, wires in ordinary circumstances) behave roughly as if Maxwell's laws were true of them.

The premise **(P0″)** looks like it might be true, and insofar as it does not permit the application of Maxwell's equations to isolated electrons, it might appear to avoid the problems just mentioned. But it does not really avoid such problems. A wire does not even roughly behave as if Maxwell's laws were true of it, insofar as it consists of stable matter, which Maxwell's equations cannot accommodate. Again, if by 'the wire' one means a falsified description of the wire (i.e., with the idealized ρ and $\vec{J}$ one would find in a typical textbook), then the wire *does* roughly obey Maxwell's equations. But the actual wire, even in ordinary lab conditions, does not even approximately obey Maxwell's equations.

Let us consider a third attempt. In the previous attempt, we tried to restrict Maxwell's equations to specific (macroscopic) objects in ordinary circumstances, to no avail. We might instead try to restrict the application of Maxwell's equations to certain *regimes* – that is, we might claim that Maxwell's equations are to be applied only when all currents, charges, energies, lengths (etc.) lie within a certain range. This will enable us to apply Maxwell's equations to the conduction current running through the wire, but will prohibit us from applying Maxwell's equations to the circulating current of any bound electron about its associated nucleus. Thus, we replace **(P0″)** with **(P0‴)**:

(P0‴) In certain specified regimes, we may assume that the E and B fields generated by certain currents and charges obey Maxwell's equations.

Can premises **(P0‴)** and **(P1′)** serve as the basis for an argument with only true premises, yielding the desired conclusion?

Again, the answer is no. The problem is that **(P0‴)** tells us nothing about what happens *outside* the relevant regimes, and therefore gives us no basis on which to assume that the overall behavior of the wires will be primarily determined by the currents, charges (etc.) within the regime of interest. We cannot know, for instance, that what happens at the atomic level can effectively be ignored. The situation is similar to someone who wants to know whether it is safe to place his hand in a bag and grab a marble. He knows, let us say, that there are red marbles and blue marbles, and he knows that the red marbles are safe to touch. However, this information, on its own, cannot form the basis for a persuasive argument that it is safe to grab a marble – perhaps the blue marbles are sizzling hot. In the same way, **(P0‴)** does not give us enough information with which to conclude that our naive Ampere's law calculation is trustworthy. We know that we can apply Maxwell's equations to the conduction current; but we have no way of knowing whether the effects of bound currents can be ignored. So this attempt at sanitizing also fails.

One might try to get around this problem by simply asserting that effects from other regimes turn out to be negligible, in ordinary lab circumstances. This will be our fourth attempt; we simply declare:

(P0⁗) In certain circumstances, the behavior of an object can be approximated by working with the $\rho, \vec{J}$ from a specified regime, and ignoring all other electromagnetic contributions.

Thus, when considering the wires, we can take $\rho = 0$ and assume $\vec{J}$ is given by the straightforward conduction current, ignoring any effects from microscopic currents.

But **(P0'''')** is also false. If $\rho = 0$ everywhere and $\vec{J} = j\hat{z}$ along the wire, the wire stops behaving like a normal solid, and assumes a totally malleable form, that responds to bumps, forces and irregularities in unusual ways, and that lacks the cohesive properties found in ordinary wires. Nature maintains the solid state of the wire through microscopic electromagnetic forces between molecules that help preserve the solid state of the wires. But such forces are precisely the forces we need to ignore, in order to avoid the problems we encountered in the previous attempts! So we *cannot* completely ignore microscopic contributions, and yet, as our previous attempts show, we cannot accommodate them (except in the most contrived situations) – unless we are happy to use an idealized description of reality.[3]

The lesson we must draw from all these failed attempts is that the simple calculation using Ampere's laws and the wires is a very complicated thing indeed, if one wants to only tell the truth. The world does *not* behave approximately as if Maxwell's laws were true in any natural or useful sense. Instead, even in simple cases, we must make false assumptions (that the wire is straight, and/or that Maxwell's equations are literally true), and use those false assumptions *very selectively*, in order to determine the behavior of nature. The calculations of classical electrodynamics therefore provide us with a good example of how bodies of knowledge in physics are *not* closed under arbitrary mathematical consequence, but must rather be used inferentially restrictively instead – thereby providing us with an important piece of evidence for the Principle of Inferential Restrictiveness and the Deduction Failure Thesis. Specifically, insofar as bodies of knowledge in physics often contain idealized premises, they cannot be closed under arbitrary mathematical inferences. For instance, from the idealized assumption that 'the wire is straight' it follows that 'the wire has no curvature', which a cursory examination will reveal to be false. All sorts of empirically false assertions follow from idealized assumptions, which is why bodies of knowledge containing idealizations are not generally closed under arbitrary mathematical inferences.

One might wonder whether the problems of this section are peculiar to classical electro-

[3]Variations on this attempt fail for similar reasons. One might want to insist that $\vec{J}$ is 'glued' into place, and so acts something like a boundary condition which remains preserved through the calculation. However, if as in our present case we want to study the *motion* of the wire, this will obviously not do.

dynamics. I do not think that they are – in fact, I think that in other areas the problems are even more acute. For instance, what does it even *mean* to apply classical mechanics to a true description of reality, when reality is quantum in nature? Difficulties in answering this question suggest that the use of classical mechanics in describing reality must involve careful lies as well. I think a systematic study of the way in which idealizations are used in physics reveals many of the same problems we have just seen in the case of classical electrodynamics. Specifically, I think the only way to make sense of how physicists reason with idealizations is within the context of an inferentially restrictive methodology. In order to get to other issues, however, we shall have to make do with the case study presented.

I shall develop some of these points further in the remainder of the paper, but before I do this, I shall consider one more slightly different attempt at sanitizing.

2.2.4 One more false start.

The fifth (and final) attempt as sanitizing is simpler, yet in some ways more successful than the previous attempts.

Let us think about how Ampere's law might have been discovered. Specifically, let us imagine Ampere observing a connection between current-measurements and magnetic field-measurements, and wanting to codify these observations into a law. In such a situation, there are (at least) two sorts of laws that could be posited. If he wanted to stick to the 'observed facts', Ampere might well have considered the following sort of law:

(L1) In ordinary circumstances, when the reading on one's ammeter is I (to a couple of significant figures), one will measure a magnetic force $\vec{B}$ (to a couple of significant figures) related to I via the equation $\oint \vec{B}.d\vec{l} = \mu_0 I$, where the line integral is performed over a closed contour surrounding the wire.

This law concerns only relationships between readings of measurement instruments. Alternatively, he might (and did) state the law in a more abstract way that does not mention readings or measurements of any sort:

(L2) $\oint \vec{B}.d\vec{l} = \mu_0 I$, where I is the current, $\vec{B}$ the magnetic field, and the line integral is performed over a closed contour surrounding the wire.[4]

Now, law **(L1)** *can* be used in a persuasive argument containing only true premises and having the desired conclusion, in a fairly straightforward way – the physicist simply notes that his ammeter readings have a particular value, and then uses **(L1)** to determine, with a bit of mathematics, what sort of attraction he should observe in the wires, to a couple of significant figures. As long one is mindful enough about the 'in ordinary circumstances' proviso, **(L1)** and all the premises employed in the subsequent argument are true.

The problem, however, in doing all this is that we violate our intuition that the physicist should be able to justify his conclusions not merely by citing observed facts about ammeters and the attraction between wires, but rather by citing laws of nature like **(L2)**, which are supposed to *explain* such facts about ammeters and wires. We are left unable to view the physicist as someone who is able to persuade himself that it is because of Ampere's law, in the form **(L2)** – and not just some observed regularity about wires and ammeters, such as **(L1)** – that the wires should attract each other with some given magnitude, to a significant figure or two. If a physicist is happy to be operationalistic enough about his discipline, he can use laws like **(L1)** to get his conclusions, avoiding false assumptions altogether. But if he wants to use laws like **(L2)** to justify his calculations, as I think he does, then he is stuck with having to be persuaded by arguments with premises he knows to be false.[5] He is therefore forced to reason inferentially restrictively, in accordance with the Principle of Inferential Restrictiveness and the Deduction Failure Thesis.

So on a view about physics which is operationalistic enough, the problem with which I am concerned can be avoided, and our physicist's argument *can* be sanitized. I proceed, however, from the assumption that this sort of operationalism is unattractive. Given that my main point lies elsewhere, I do not wish to spend time going over the problems with

[4]An instrumentalist might be inclined to take **(L2)** to simply *mean* something like **(L1)**. However, I intend **(L2)** to be a statement about physical reality independent of facts about measuring devices. It is therefore the sort of claim the instrumentalist would deny, or perhaps even find meaningless.

[5]One might worry that insofar as **(L1)** is supposed to be derivable from **(L2)**, any argument with **(L1)** as a premise can be replaced by one with **(L2)** as a premise. Using the fact that there are arguments involving only **(L1)** that solve our problem, it then follows that there are arguments involving **(L2)** that also solve our problem. But there are two difficulties. First, **(L2)** is false, so the new argument contains a false premise, and does not solve our problem. Second, any theoretical derivation of **(L1)** from **(L2)** will have to involve idealizations in the way we represent the workings of our ammeters and other measurement equipment, forcing our original problem to bite us in the tail.

operationalism, many of which have been discussed already. (See, for example, [44], [69] and [29].) Operationalism of a certain sort offers a way around my problem; but the price, in my view, is too high to pay – especially when other solutions are available, as we shall see.

2.3 INFERENTIAL RESTRICTIVENESS

2.3.1 Inferential Restrictiveness as a Solution.

I have argued that many persuasive arguments in physics should be viewed as having false premises. The most natural way to think of what happens when our physicist does his Ampere's law calculation, then, is that he invokes some false description of reality, but in such a way that his conclusion remains trustworthy. Let us think a bit about how the physicist does this.

When the physicist with the wires begins his argument by telling us that the wires are infinitely thin line segments of uniform charge, he is assuming that he will only be required to calculate the first couple of significant figures of the force of attraction between the wires. He assumes (correctly) that at no later point will he be asked about the 3rd, 4th, or 20th significant figure of any associated quantity. The physicist fully understands that *false* claims can be inferred from the claim that the wires are infinitely thin line segments of uniform charge; but given his limited inferential aims – namely, making some predictions about the rough values of certain simple quantities – he takes his false description of the wires to be trustworthy. Likewise, when he invokes Maxwell's equations, he assumes (correctly) that he will not be quizzed about the stability of the atoms making up the wire – he knows in advance that he will only use Maxwell's equations for a very simple purpose, and that they are therefore trustworthy. Given modest enough goals, arguments containing highly idealized (and therefore false) premises can be trusted; but one cannot trust arguments containing such premises if one is trying to draw especially detailed conclusions. As the need for detail increases, so must the subtlety of one's lies.

The central idea here is that a premise known to be false can sometimes be used with confidence if one plans on involving it only in a certain form of argument. In our example, the physicist may make the assumption that the wire is an infinitely thin line segment, and draw conclusions concerning the first two significant figures of certain physical quantities; but he may not draw any conclusions about other things, such as the 3rd, 4th or 20th significant figures of those quantities. Similar things can be said about his use of Maxwell's laws. As long as he restricts the inferences in which his false premises are involved, he can avoid disaster.

Another way of putting this is to say that the physicist's methodology is *inferentially restrictive*. When the physicist makes some sort of idealizing assumption about a physical system (as he must, for there is no system about which he has complete information) there are always some inferences involving the idealizing assumption that ought not be trusted. There will always be restrictions on the inferential use to which one can put an idealized assumption. Arguments that violate the (often implicit) inferential restrictions that accompany an idealized assumption cannot be regarded as persuasive, while arguments that respect the inferential restrictions can. Associated with any idealized assumption is a set of inferential restrictions that determine which arguments involving it should be regarded as persuasive and which should not. Insofar as idealized assumptions are an indispensable part of applying physics to reality, so too are inferential restrictions. Note that the presence of inferential restrictions provided evidence for the Principle of Inferential Restrictiveness; the necessary use of inferential restrictions provides evidence for the Deduction Failure Principle.

The problem of understanding how false premises can be used in persuasive arguments is therefore resolved by pointing out that false premises can still be inferentially useful when accompanied by a corresponding set of inferential restrictions. This is what happens when our physicist asserts, for instance, that the wires before him are straight – an assertion he knows to be false – and then infers facts that he takes to be trustworthy.

One might worry, however, that there is some sort of double standard in play here. In the previous section, I argued that it is impossible to reconstruct a certain argument in such a way that (i) the argument has the desired conclusion, (ii) the argument has only true premises, and (iii) one is *justified* in asserting the conclusion, given the premises. But does inferential restrictiveness do any better? Inferential restrictiveness allows us to ignore some of the problems that bogged us down in the previous section, but does it do better on count (iii)? In what sense are we justified in believing the conclusion of an argument whose premises may well even be *known* to be false?

There are a couple of different answers to this question. When one introduces inferential restrictions, conclusions of arguments are still required to follow (logically) from their premises. In this sense, we are perfectly justified in believing the conclusion of such an argument, given its premises.

However, the problem is that *false* conclusions also follow from the premises in question. So the main question here is: can we *ever* be justified in believing the conclusion of an argument with premises that we know to be false? If so, how? We will address this question briefly in the next section (section 3.2), but more fully in Section 4.

2.3.2 Where Do Inferential Restrictions Come From?

In order to begin to answer the question posed at the end of the last section, it will be helpful to consider where the set of inferential restrictions associated with an idealizing assumption comes from. Given his goals, how does the physicist know to what sorts of uses a false description of reality can or cannot be put? One might suspect that the only thing that can answer these questions is something like perturbation theory. But we know that this cannot be the correct answer, insofar as even the use of perturbation theory presupposes false premises.

So if perturbation theory is not the source of the physicist's knowledge of what sorts of falsehoods can be put to use in persuasive arguments, what is? I shall argue that this sort of knowledge has its source in *experience* and *conjecture*, rather than theoretical deductions of any sort. I shall also discuss this idea in Chapter 4.

When physicists make predictions, what they are generally doing is applying some set of laws to an idealized (and hence false) description of some physical system. A-priori, there is no guarantee that anything but nonsense will be produced. All the physicist can do is give it a shot, and see if anything useful happens. In many cases, nonsense *is* produced; many idealized ways of describing physical systems turn out to not be very useful. If one tries to treat a human being as a point mass being pushed around by the forces of Newtonian mechanics, one does not do very well – but if we treat the solar system in this way, we do excellently. This is not something that could be ascertained in advance. It is by no means a conceptual truth that people cannot be usefully treated as point masses, or that planets can.

One learns by experience that in some inferential contexts, certain types of lies work well, and that in other contexts, different lies are necessary. After enough experience, the physicist begins to formulate heuristic principles about what sorts of lies work well in what contexts – for example, that sets of interacting, distant bodies can often be treated as classical point masses provided only modest conclusions are drawn, or that in a certain 'regime', particles may be treated in a certain way. The history of physics is very much the process of sniffing out ways in which false descriptions of reality can be put to use, and exploring the limits to which these lies can be pushed. There are, however, no a-priori truths in the art of lying; one must simply see what one can get away with by trying, or by watching the failures of others.

Heuristic principles about which lies are useful and which inferential restrictions are necessary will rarely be precisely formulated; they will often be vague, ambiguous, and sometimes even conflict with each other. (Is it better to treat a highly viscous fluid as a thermodynamic ensemble, or a classical body? In borderline cases, our principles do not give a decisive answer.) The *conjectural* nature of these heuristic principles must also be emphasized – like any set of principles drawn from experience, they will often need to be revised, or even discarded, in light of subsequent experience. The physicist can rarely be *completely* sure that an inference from some false set of premises is harmless; at best, the trustworthiness of such an inference may simply be a good conjecture. Again, we will comment more on this in Chapter 4.

In spite of all this, it would be incorrect to claim that theoretical considerations play *no* role whatsoever in figuring out what sorts of inferences are appropriate, given some false description of reality. One might learn by experience, for instance, that classical electrodynamics is trustworthy in a certain regime, at which stage one can use the perturbation theory of classical electrodynamics to deduce the effect of some approximation in that situation. The judgment that electrodynamic perturbation theory is appropriate in a certain regime implicitly codifies much antecedent experience about which sorts of lies concerning electromagnetism can be trusted in which sorts of contexts, and implicitly draws on protracted experimental struggles which had to be fought hard and won. Indeed, the succinct summarization of all these experimental struggles into a single rule of thumb – trust electromagnetic perturbation theory in such and such a regime – is an extraordinary *experimental* achievement, that provides the physicist with a uniform method for choosing his inferential restrictions when dealing with electromagnetic phenomena. In such a case, theoretical deduction plays a *significant* role in guiding the physicist's sense of the limits to which lies can be pushed. But still, this use of theory is impossible were it not for the judgment that classical electrodynamics is appropriate in a given regime – a judgment that can come only from experience. So, while theoretical calculations can play a role in justifying inferential restrictions, they do so only because *other*, more fundamental, ways of lying have been justified by *experience*.

Moreover, the discovery of a physical law, and an understanding of how one must lie if one is to apply that law, often come together. It is not as if Ampere conjectured his law, verified his law, and then tackled the problem of finding useful ways of lying about physical situations in order to apply his law. Rather, at the stage of trying to verify his law, Ampere already had to decide what sort inferences involving his law could be trusted, given some false description of reality. Only then could any empirical 'verification' occur. Ampere proposed his law, together with some idealized representation of a physical situation to which he wanted to apply his law, conjectured that the subsequent inferences should be trusted, and then put it all to the test. In verifying his law, he was therefore *already* tackling the question of which ways of lying are to be trusted when his law was being used. Indeed, the physicist never verifies a law on its own; what he verifies is that the law, plus some way of lying (which includes a set of inferential restrictions) is useful. The discovery of laws, and the development of inferential restrictions that are useful when applying that law, often occur in one and the same breath.

All this helps us to address the question of how we can be justified in believing the conclusion of an argument with premises that we know to be false. Our laws are designed in such a way that their application presupposes some element of lying about the world. To say that the law has been verified is therefore just to say that the drawing of certain inferences from certain false assumptions has been shown to be trustworthy. We can trust certain inferences from false premises because a mass of experience tells us so. Indeed, the very formulation of mathematical laws of nature presupposes that we have learned which sorts of lies are trustworthy in which sorts of contexts. Thus, there is a great deal of inductive evidence which helps to justify the use of idealizations. (We will pick up on some related themes in Chapter 4.)

However, a critic might remind us that to say a mode of argument works does not *justify* that mode of argument. (The literature on the problem of induction is a nice illustration of this point.) We will therefore need to say more about whether inferential restrictiveness can ever be *justified* in this stronger sense. We will pick this issue back up shortly, after a brief digression.

2.4 DO THE LAWS OF PHYSICS LIE?

In this section, I wish to make a brief digression, and discuss some of the claims that Cartwright makes in her controversial [17].

In her book [17], Cartwright makes the striking claim that *'... the fundamental equations [of physics] do not govern objects in reality ...'* (p 129). Specifically, she claims that:

> *'... it usually does not make sense to talk of the fundamental laws of nature playing out their consequences in reality. For the kind of antecedent situations that fall under the fundamental laws are generally the fictional situations of a model, prepared for the needs of the theory, and not the blousy situations of reality.'*

(See p. 160.) Her argument is as follows: only when we 'idealize' real physical systems, i.e., only when we strip real physical systems of much of their messiness, and substitute smoothness and order instead, can we can tell whether or not a physical systems obeys the fundamental laws of physics. But if this is right, then at the end of the day all we have the right to say is that the fundamental laws of physics hold in situations somewhat similar to those we find in the actual world, but neatened up around the edges. All we can say is that the laws of physics hold in certain *fictional* situations. (On page 129, Cartwright makes the point by telling us that *'... fundamental equations ... govern only objects in models.'*)[6] Cartwright furthermore suggests that insofar as we want to explain facts about the real, unidealized world, we must turn to ceteris paribus laws, but this later thesis will not concern us here.

I am (obviously) sympathetic with Cartwright's claim that it is only once we idealize real physical systems that we can tell whether or not they obey the laws of physics. In our example, our physicist must pretend that his wires are straight before he can put Ampere's law to the test. Even if he decides to be more subtle, and account for some curvature in the wires, he must idealize *somewhere* in his description of the physical system. But does it follow from this that we only have the right to say that Ampere's law has been confirmed in a *fictional* situation?

Let us use some of the concepts introduced in the previous sections to tackle this question. Cartwright thinks that, when we strip a real physical system of much of its messiness, and substitute smoothness and order instead, we move from a description of a real physical system to a description of a fictional system. While this may be true sometimes, we must be careful. We can sometimes use a falsehood to describe a system in a useful way as long as we impose inferential restrictions. As long as these inferential restrictions are respected, any conclusions drawn using the false premise can be said to hold of the original, *unidealized* system. The phenomenon of inferential restrictiveness shows that one can reason with a fiction, and still draw conclusions about a *real* system.

Thus, when the physicist in our example imagines that the wires before him are straight, the conclusions he draws from this may be said to hold of the original, physical system, as long as he obeys the relevant inferential restrictions. In obeying these inferential restrictions, he manages to draw a conclusion about the actual wires before him, and not about some fictional pair of wires similar to those before him. Thus, we have the right to say that we have confirmed that Ampere's law holds *in the real world.*

So the laws of physics do not lie. It is certainly true that we must lie when applying the laws of physics. But if one is careful with one's lies, one's conclusions can still be true.

[6] Over the course of her book, Cartwright presents several arguments for this main conclusion. The one I have summarized, however, is perhaps the most compelling, and is the only one I shall give any attention to.

Physics does not degenerate into theater (as Cartwright is fond of saying) – it simply uses a bit of theater on the way. Cannot good theater, after all, sometimes teach us something useful about the real world?

2.5 JUSTIFYING INFERENTIAL RESTRICTIVENESS.

2.5.1 Can Inferential Restrictiveness be Justified?

In this section, I wish to focus exclusively on the question of how we can be justified in believing the conclusion of an argument with false premises.

Why are physicists so comfortable reasoning with false premises? One might reply that they are comfortable doing so because *they have to* do so; the world is so complicated, and our understanding of it so incomplete, that the physicist has no choice but to reason with simplified falsehoods.

But this answer is not completely convincing. Faced with the possibility of uttering a falsehood, the scientist can always choose to remain silent, and pursue different avenues of investigation – avenues of investigation perhaps reminiscent of some of the less mathematical sciences, where the problems I have been discussing do not appear so acutely. But the physicist does not pursue this route – why not?

One might reply that the physicist chooses not to remain silent, because reasoning with false premises turns out to be so effective. If something works well, it is surely worth pursuing, even if it involves a behavior that one might not otherwise endorse.

But none of this brings us any closer to understanding *how it is* that we can reason successfully with false premises. We have discovered that reasoning with false premises works – but, on its own, this is no justification for doing so, just as empirically discovering that adding the digits of a number produces a test for divisibility by 9 does not count as a *justification* for this procedure.

Let us consider one way in which we might try to justify inferential restrictiveness. Recall the fact that, in arguing that the pair of wires we considered earlier were attracted in some particular way, we were forced to use an argument with false premises. All attempts to sanitize the argument – that is, to replace it with an argument with true premises and the same conclusion – were found to fail. Some of this failure was due to the fact that our understanding of the laws of physics was incomplete. Thus, one might hope that, given a complete understanding of the ultimate laws of physics, the argument in question *could* be sanitized. Consider the following suggestion: to embrace an inferentially restrictive methodology is just to endorse a set of arguments which, although not presently sanitizable, one conjectures *will* ultimately be sanitizable, once the final laws of physics are known. If this conjecture turns out to be true, inferential restrictiveness is justifiable, insofar as arguments respecting the relevant inferential restrictions *can* be sanitized.

I do not think, however, that this is a very satisfying justification of inferential restrictiveness. The main problem is that all we get is that inferential restrictiveness *might* be justified. It might turn out, however, that the ultimate laws of nature are such that they do not admit of anything like standard perturbation theory – or it might turn out that standard perturbation theory does not yield the conclusions we want. Moreover, it seems impossible to argue that we have any sort of inductive evidence to expect things to work out one way or another, so it is not as if we can even argue that the relevant conjecture, while just a conjecture, is nevertheless a *reasonable* one.

Of course, one possibility is that the argument given is simply the best that can be done. Short of deriving facts from the ultimate laws of nature, perhaps the best the physicist can do is to give arguments that *might* turn out to be justified. Perhaps demanding anything more is inappropriate. If this is right, inferential restrictiveness gives us everything that can be humanly obtained, and to ask for more is unreasonable.

However, I do not think we should be so quick to embrace such a conclusion. I think that there are other ways in which inferential restrictiveness can be justified, as we shall see.

2.5.2 Flashing Lights.

Let us consider in a little more detail the question of how to justify inferential restrictiveness.

Roughly speaking, I would like to argue that it is possible for the world to be built in such a way that, while certain sorts of mathematical inferences from certain premises are reliable, other sorts of (otherwise valid) mathematical inferences from those same premises might *not* be reliable. I would like to think about this property of the world as a 'fact' about the world (in much the same way that ordinary laws are facts about the world) that can be used to justify inferential restrictiveness. It will be helpful to spell out in considerable detail exactly what sort of 'fact' about the world we could be dealing with here. It will be simplest to do this with a toy model.

Before going into the details of the toy model, however, it is worth noting that something quite similar should be familiar from a different context. It is a 'fact' about the world that certain classical calculations are quite accurate for speeds $\ll c$, even though other classical calculations – in the same or different contexts – can lead us badly astray. Thus, beginning with the laws of Newtonian mechanics, certain sorts of purely mathematical arguments are reliable, while other purely mathematical arguments are not. Such facts about what sorts of mathematical arguments are trustworthy, given the laws of Newtonian mechanics as premises, are 'facts' about the world in much the same way that *any* law is a 'fact' about the world. This is the kind of phenomenon that I would like to treat in more rigorous detail here.

Let us turn to the toy model. Imagine a system of three lights, A, B, and C. Each of the three lights occasionally flashes on for a very short period of time. The following observations are made:

(O1) Each individual light has a period τ such that the light seems to flash on roughly every τ seconds.

(O2) Whenever light A flashes on, light B or C also flashes on.

(O3) Lights A and B have been observed to simultaneously flash on, with light C remaining off.

(O4) Lights A and C have been observed to simultaneously flash on, with light B remaining off.

(O5) All three lights have been observed to simultaneously flash on.

Here is one possible scenario in which all 5 of these claims are true: imagine that at time $t = 0$, all the lights flash on, that light B flashes on every .002 seconds afterwards, light C flashes on every .003 seconds afterwards, and that light A flashes on at times $t =$

$0, 0.999, 2, 3, 3.999, 5, 6, 6.999, ...$ Note that **(O1)** is true, because it only demands that the flashes are *roughly* periodic.

Let us imagine, however, that, it has not yet been realized that the word 'roughly' in **(O1)** must be taken seriously. Let us imagine then that a scientist tries to construct a theory accommodating **(O1)**–**(O5)** in the following way: it is assumed that at $t = 0$ all three lights flash on, and it is assumed that there are positive real numbers τ_A, τ_B, τ_C such that light A flashes on at integral multiples of τ_A, light B at integral multiples of τ_B, and light C at integral multiples of τ_C.

It is easily seen that such a theory cannot be built. For at time $t = \tau_A$, by **(O2)**, either light B or C must be on. Let us assume that it is light B; then we must have $\tau_B = (\tau_A)/n$ for some natural number n. From this it follows that whenever light A is on, light B will be on, violating **(O4)**.

There are various things the physicist can do in this situation. I wish to consider just one possibility. Imagine the physicist does some careful measurements, decides that $\tau_A = 1s, \tau_B = 0.002s$, and $\tau_C = 0.003s$. (Even though he is, of course, technically incorrect in saying that $\tau_A = 1s$, his claim is quite accurate – the time of As actual flashing will deviate from our physicist's predictions by less than 1 percent.) He recognizes, of course, that at and about $t = 1, 4, 7, ...$, this leads to a divergence with what has been observed in the lab. In order to avoid a contradiction with known experimental results, he therefore also declares that only very specific conclusions concerning times $t = 1, 4, 7, ...$ and their immediate vicinity should be drawn from his assumptions. In particular, one can assume that in the vicinity of $t = 1, 4, 7, ...$, light A will flash once, and light B or C with it, but one cannot make any assumptions about whether it will be B or C that flashes with A, or precisely when. (We can assume, however, that at $t = 0$, all three lights flash.) Away from $t = 1, 4, 7, ...$, however, calculations based on his values of τ_A, τ_B, and τ_C are to be trusted.

What the physicist has just done, of course, is to introduce an inferential restriction. He has given us a set of basic assumptions that can be used to deduce further facts about the three light system. However, he has also identified a family of inferences which he recognizes as unsafe. In such cases, he simply says that his premises are being taken too literally, that one is letting the mathematics lead one astray, or some such thing.

Having done all this, we then imagine that our physicist produces the following argument:

(Q1) Lights A and B both flash on at $t = 0$,

(Q2) Light A has a period of $1s$,

(Q3) Light B has a period of $0.002s$,

(Q4) Light A is flashing now,

(C) Therefore, light B will flash at least 500 times (approximately) before light A flashes again.[7]

This argument works nicely, and its conclusion is true. The problem, however, is that the conjunction of premises **(Q1)**, **(Q2)** and **(Q3)** are *false*, insofar as they entail that whenever light A is on, light B is also on – a fact that is experimentally known to be incorrect. Nevertheless, by introducing an inferential restriction, our philosopher can reason quite successfully with **(Q1)**, **(Q2)** and **(Q3)**. So our main question, applied to this simpler case, now becomes – why, in the previous argument, is our philosopher justified in asserting the conclusion **(C)**, given the falsehood of the premises **(Q1)**, **(Q2)** and **(Q3)**? Is the fact that the conclusion **(C)** turns out to be true just a co-incidence? Is it the result of some sort of gerrymandering, or data-mining? Or is the conclusion in some sense *justified* by the (false) premises? I shall argue for this last possibility.

In order to do so, let us consider some definitions. We may say of the three light system that its behavior is determined by a fundamental set of laws L, where L consists of the claims that at $t = 0$, all the lights flash, that $\tau_B = .002s, \tau_C = .003s$, and A flashes at times $t = 0, 0.999, 2, 3, 3.999, 5, 6, 6.999, ...$ This set of laws is, of course, not (completely) known to the physicist.

Let T_1 be the set of claims that are consistent with L, and that might be made of the three light system. Typical elements of T_1 might be: 'both A and B are now flashing', 'between every two flashes of A there are flashes of B and C', 'A will flash within the next $0.3s$, and at that time, C will also flash', 'the flashes of A are not exactly periodic', and so

[7] The physicist cannot conclude that it will flash *exactly* 500 times, because his inferential restrictions demand that he profess ignorance about when and how much B flashes in the immediate vicinity of $t = 1, 4, 7, ...$

on. These are judgments that we can imagine God making (where the only divine attribute I need to exploit is that God knows L.)

For sentences $p, q \in T_1$, say that $p \leq_1 q$ iff q may be deduced from p and L. So, for instance, if p_1 = 'lights B and C have just flashed 100 times simultaneously without light A flashing at all', and q_1 = 'light A will flash sometime within the next .5 seconds', then $p_1 \leq_1 q_1$; if p_2 ='the last time light A flashed, lights B and C also flashed', and q_2 = 'the next time light A flashes, light C will also flash', then $p_2 \leq_1 q_2$. For $p, q \in T_1$, define $p =_1 q$ iff $p \leq_1 q$ and $q \leq_1 p$. So for instance, if p = 'A is flashing and B is not flashing', and q = 'A is flashing, B is not flashing, and C is flashing', then $q =_1 p$.

For any $p \in T_1$, let $[p]_1$ be the set of $q \in T_1$ such that $p =_1 q$. For $p, q \in T_1$, define $[p]_1 \leq_{[1]} [q]_1$ iff $p \leq_1 q$.[8] Then $\leq_{[1]}$ defines a partial order on the set of equivalence classes $[p]_1$ with $p \in T_1$; call the resulting partially ordered set P_1. The partially ordered set P_1 encodes the implications that hold between assertions about the world from God's perspective, i.e., in virtue of the physical laws being as they are.

Now let T_2 be the set of claims that could be made about the world, and that are consistent with the 4 claims:

(R1) Lights A, B, and C all flash on at $t = 0$,

(R2) Light A has a period of $1s$,

(R3) Light B has a period of $0.002s$,

(R4) Light C has a period of $0.003s$.

For instance, T_2 contains 'whenever A flashes, B flashes', 'A, B and C are now flashing', and 'B and C will flash together at least 15 more times before A flashes.'

For $p, q \in T_2$, define $p \leq_2 q$ iff q is mathematically deducible from p, subject to the inferential restriction that the deduction rely on no facts about times within ± 0.002 of $t = 1, 4, 7, ...$, except that A will flash at one point in each such interval, and with it B and/or C. So, for instance, if p_1 = 'light A has just flashed', and q_1 = 'either light B or C will flash next', then $p_1 \leq_2 q_1$; but if p_2 = 'light A has just flashed with B', and q_2 = 'the next time light A flashes, light B will also flash', then $p_2 \not\leq_2 q_2$. As before, define

[8] Note that if $p_1 =_1 p_2$ and $q_1 =_1 q_2$ then $p_1 \leq_1 q_1$ iff $p_1 \leq_1 q_1$, so this definition makes sense.

$p =_2 q$ iff $p \leq_2 q$ and $q \leq_2 p$. Also define equivalence classes $[p]_2$ for $p \in T_2$ and an induced partial ordering $\leq_{[2]}$ on these equivalence classes as before, and call the resulting partially ordered set P_2. The partially ordered set P_2 encodes the mathematical implications that hold between assertions about the world, in virtue of the *idealized* assumptions **(R1)–(R4)** that have been made, and subject to the inferential restrictions just mentioned.

There are many propositions in both T_1 and T_2 – for instance, 'lights A and B both flash on at $t = 0$', 'lights B and C will flash simultaneously at least 20 more times before A flashes', and so on. Given any such $p, q \in (T_1 \cap T_2)$, we can order p and q with either $\leq_1$ or $\leq_2$.[9] It turns out, however, that there is a type of 'agreement' between these two different orderings, as the following result shows:

Poset Homomorphism Result: $\forall p, q \in (T_1 \cap T_2)[([p]_2 \leq_{[2]} [q]_2) \rightarrow ([p]_1 \leq_{[1]} [q]_1)]$.

This result holds, because our inferential restrictions are designed so that any argument that is valid under the assumptions **(R1)–(R4)**, and respects all inferential restrictions, will also be valid when **(R1)–(R4)** are replaced by their counterparts in L. In other words, we have a uniform strategy for sanitizing those arguments with premises **(R1)–(R4)** that respect all relevant inferential restrictions. (Our physicist, however, cannot do the sanitizing himself, because he does not yet know the contents of L.)

Note that the implication '$\rightarrow$' in the Poset Homomorphism Result cannot be replaced by a biconditional '$\leftrightarrow$'. For instance, if $p =$ 'A, B, and C just flashed simultaneously' and $q =$ 'the next time A flashes, C will flash', then $[p]_1 \leq_1 [q]_1$, but $[p]_2 \not\leq_2 [q]_2$.

What does the Poset Homomorphism Result tell us? Technically, it tells us that there is a homomorphic map of the partially ordered set P_2 into P_1. That is, it tells us that the $[\ ]_2$ equivalence classes are contained in $[\ ]_1$ equivalence classes, and thus that we can define a function F from the set of $[\ ]_2$ equivalence classes into the set of $[\ ]_1$ equivalence classes, such that for all $[\ ]_2$ equivalence classes x and y, $x \leq_2 y \rightarrow F[x] \leq_1 F[y]$. Such a mapping F is a 'homomorphism' of relational structures.

What this means is that any simple mathematical argument in the *idealized* system (i.e., any argument using the premises **(R1)–(R4)** that respects all inferential restrictions), yields

[9]Strictly speaking, we are not ordering p and q, but $[p]_1$ and $[q]_1$, and $[p]_2$ and $[q]_2$; this technicality does not affect anything that follows.

statements that are true in virtue of the *unidealized* (and, to our physicist, *unknown*) laws L. Our inferentially restricted system therefore 'tracks the truth' about the world, in spite of its use of false premises. Specifically, we have defined a class of statements $T = (T_1 \cap T_2)$ such that the mathematical implications that hold between elements of T, in our idealized, inferentially restricted system of reasoning, are mirrored by implications that hold between such elements of T in virtue of the *actual laws* of nature being as they are. Our idealized, inferentially restricted system of 'mirrors' actuality in this way.

However, not *all* relations between elements of T that hold in virtue of the actual laws of nature being as they are will be mirrored by corresponding relations that hold in our idealized, inferentially restricted system of reasoning. This is to be expected – when we idealize, we gain a certain ease in moving between certain facts about reality, but at a cost; very detailed facts about reality lie beyond our grasp. It is enough to ask that the relations that hold between elements of T in our idealized, inferentially restricted system of reasoning be mirrored by relations that hold in virtue of the *actual laws* of nature being as they are – to ask for the converse, however, would be to ask for too much.[10]

These observations help us to produce a justification of inferential restrictiveness in the three light system. The Poset Homomorphism result tells us that chains of correct reasoning in our idealized, inferentially restricted system mirror correct chains of reasoning in the unidealized system. It is because of this result that we are *justified* in believing the conclusion **(C)** of the argument with premises **(Q1)**–**(Q4)** above. The Poset Homomorphism result therefore justifies the conclusion of certain arguments based on certain idealized (i.e., literally false) assumptions, provided that inferential restrictions are obeyed.

One can imagine a couple of objections to my claim that the Poset Homomorphism Result can justify our belief in the conclusion of an argument with false premises. The first objection is that the Poset Homomorphism Result simply repeats the claim that inferential restrictiveness 'works'. (It repeats it in fancier language, but that should not fool us). But the claim that inferential restrictiveness is justified must in some way go beyond the mere claim that it works. Insofar as the Poset Homomorphism result does not do this, it fails to

[10] Of course, this point is related to the fact that we cannot replace the '$\rightarrow$' by a '$\leftrightarrow$' in the Poset Homomorphism Result.

count as a justification of the sort desired.

But I do not think that this objection is effective. The Poset Homomorphism Result functions as a general law, in terms of which the conclusions of particular inferences may be justified. In order to make this point clear, it will help to consider an analogy.

Imagine that a physicist applies Schrödinger's equation to a system, and makes some prediction about the value of a future measurement. Imagine that the physicist is then challenged to explain how Schrödinger's equation justifies the assertion that the measurement will turn out in the way predicted. Let us suppose the physicist goes through all the relevant calculations, in order to answer the challenge. Imagine that the physicist is then challenged to justify his belief in Schrödinger's equation. To this, he replies by talking about the mass of relevant empirical evidence. At this point, his interrogator interrupts – Aha! All you are doing is telling me that Schrödinger's equation *works*. But I was asking for a *justification* of your claim that the measurement in question will turn out to be such-and-such, and you still haven't really answered me. All you've managed to say is that a certain family of calculations work because they work.

Such an interrogator, however, is surely in error, in (at least) two ways. First, Schrödinger's equation surely *does* justify the assertion that the measurement will come out in such-and-such a way, independently of how one, in turn, justifies Schrödinger's equation. Second, unless Schrödinger's equation is derivable from more fundamental laws, Schrödinger's equation does not call for further justification – it is simply a true fact about the world. Of course, one can still talk about *reasons for believing* Schrödinger's equation. But from the fact that one's reasons for believing Schrödinger's equation are all inductive, it does not follow that Schrödinger's equation fails to justify the assertion that the measurement in question will have such-and-such a value.

Likewise, the Poset Homomorphism Result is a fact about the world that can be used to justify the drawing of a conclusion like **(C)** from (false) premises such as **(Q1)**–**(Q4)**. In this case, the conclusion **(C)** *is justified* in virtue of a general fact – the Poset Homomorphism Result – that relates $\leq_1$ and $\leq_2$. How exactly the Poset Homomorphism Theorem is *in turn* 'justified' (whatever this might mean), or whether even it *can be* justified (as opposed to being viewed as a 'brute fact') does not affect the status of the original justification.

Let us consider a second objection. One might argue that the justification for inferential restrictiveness in terms of the Poset Homomorphism Result, while a genuine justification, is not really a useful one. To see why, let us ask how our physicist *knows* that the Poset Homomorphism Result obtains. The worry is that it is not possible for him to know the Poset Homomorphism Result *without already knowing* that conclusions such as **(C)** (i.e., the **(C)** that follows from **(Q1)**–**(Q4)**) are true. Insofar as this is so, the proposed justification of inferential restrictiveness does not allow us to use inferential restrictive reasoning to draw conclusions that are not already known. Justifying inferential restrictiveness in terms of the Poset Homomorphism Result therefore makes it look as if it is not possible to use inferentially restrictive reasoning in a justified way to obtain new knowledge.

But I also think that this objection is unfair. The implication 'light A is flashing now' $\rightarrow$ 'light B will flash at least 500 times (approximately) before light A flashes again' need not have been known prior to the articulation of the inferentially restricted deductive system in question. (It certainly might have been, but it equally well might not have been.) Recall that it is generally through a process of experience and conjecture (discussed earlier), that our physicist arrives at his inferentially restrictive system of reasoning. The set of experiences involved in this process may or may not involve explicit verification of the implication in question (or explicit verification of facts that logically entail it.) Because inferentially restrictive systems are generally arrived at through a process of experience and conjecture, it *is* possible to use such systems of reasoning in a justified way to obtain new knowledge. This is because one can be in a position in which it is reasonable to conjecture that something like the Poset Homomorphism Result holds, without already knowing *all* the conclusions that some specific inferentially restricted system of reasoning yields. This fact is enough to answer the objection.

2.5.3 Back to Ampere's Law

The three light system discussed in the previous section is useful, but it is still only a toy model. Let us turn to the case of Ampere's law.

Let L be the true laws of nature (from God's point of view), let T_1 be the set of claims consistent with T. (We will be interested in claims such as 'two adjacent wires have approximate currents $\vec{J}_1$, $\vec{J}_2$ running through them', 'the wires are repelled', and so on.) For $p, q \in T_1$, say that $p \leq_1 q$ iff q may be deduced from p and L. Form equivalence classes $[p]_1$ as before, and call the resulting partially ordered set P_1.

Let T_2 be the set of claims consistent with classical electrodynamics. As I have shown earlier, classical electrodynamics is only useful for describing the real world if we introduce inferential restrictions – for instance, we must not interrogate the internal structure of matter too deeply, or ask too many questions about the solid state of the wire (how it coheres, responds to irregularities, bumps, and so on.) This set of inferential restrictions is a tentative and loosely articulated set of rules of thumb, unlike those associated with the three light system, but that will not affect the basic point.

For $p, q \in T_2$, let $p \leq_2 q$ iff q may be deduced from p and Maxwell's equations, in such a way that all relevant inferential restrictions are obeyed. Define equivalence classes $[p]_2$ as before, and call the resulting partially ordered set P_2.

The set $T_1 \cap T_2$ is non-empty; it contains statements such as 'these two wires behave approximately as if they have currents $\vec{J}_1$, $\vec{J}_2$ running through them', 'the wires are repelled', and so on. In addition, we have as before the following:

Poset Homomorphism Result: $\forall p, q \in (T_1 \cap T_2)[([p]_2 \leq_{[2]} [q]_2) \rightarrow ([p]_1 \leq_{[1]} [q]_1)]$.

In other words, implications between elements of $T_1 \cap T_2$ which hold in virtue of Maxwell's equations and which respect all relevant inferential restrictions, also hold in virtue of L.

The Poset Homomorphism Result is an empirical result for which a mass of empirical support can be given. Every success of classical electrodynamics is a piece of evidence in favor of the Poset Homomorphism Result. The Poset Homomorphism Result is a fact about the world – much like Schrödinger's equation – that is supported by empirical evidence, and which, in turn, can be used to justify the conclusions of many different forms of argument.

Note that the Poset Homomorphism Result is *not* best thought of as a 'Theorem' (which is why I have merely called it a 'Result'.) First, no sort of proof is possible, because we do not know L. But second, it is an empirical result, and thus not the sort of result for which we can generally demand mathematical proof, much like Schrödinger's equation. (This is not to deny, however, that a proof of the Poset Homomorphism Result *might* be possible if we knew L.)

So, with the use of the Poset Homomorphism Result, we *can* view the physicist's calculation with Ampere's law as justifying its conclusion, in spite of the fact that false premises are used along the way. The justification of conclusions obtained from idealized premises in other areas of physics will be much the same – there will be a homomorphism between sets of implications in some idealized, inferentially restricted system, and the set of implications in the axiom system L. It is because of this that – regardless of what Kant told us – lying can not only be *useful*, but *justified*. Insofar as physicists are *prepared* to lie, we find evidence for the Principle of Inferential Restrictiveness; insofar as their goals seem to *demand* that they lie, we find evidence for the Deduction Failure Principle.

3.0 IS MATHEMATICAL RIGOR NECESSARY IN PHYSICS?

3.1 INTRODUCTION.

In the last chapter, I argued that reasoning with idealizations in physics is often done with the help of inferential restrictions. In this section, I would like to examine some cases of non-rigorous mathematical reasoning in physics, and argue that they too occur only with the help of inferential restrictions.

Specifically, I would like to consider physical theories which exploit ill-defined, or even incoherent mathematical concepts (such as delta functions or path integrals in quantum mechanics). Such violations of mathematical rigor are quite common in physics, but are also quite safe – and can even be justified – so long as appropriate inferential restrictions are adopted. Thus, a certain type of violation of mathematical rigor provides us with our second main piece of evidence (I shall argue) for both the Principle of Inferential Restrictiveness, and the Deduction Failure Thesis.

3.2 MATHEMATICAL RIGOR IN PHYSICS.

I shall begin by outlining a few of the different attitudes that might be taken towards mathematical rigor in physics.

Physics textbooks are full of arguments that are unpersuasive insofar as we try to construe them as pieces of mathematics.[1] Often, the reason these arguments are unpersuasive as

[1] In order for an argument to be persuasive as a piece of mathematics, I do not require that it be written out as a proof in ZFC, or that it be rigorous by the most exacting of mathematical standards. By calling a piece of reasoning persuasive as a piece of mathematics, I simply mean that it is the sort of thing that most mathematicians would be inclined to accept as a valid piece of mathematical reasoning.

pieces of mathematics is that they violate the commonly accepted standards of mathematical rigor. Some such violations are fairly innocent, and need not bother anyone too greatly – as, for instance, when limit signs or derivatives are interchanged without the scrutiny ordinarily required by the mathematician. But other violations are not obviously so innocent. In particular, the introduction and use of mathematically ill-defined concepts (such as delta functions and path integrals in quantum mechanics) cannot be obviously dismissed as merely insignificant violations of the usual standards of mathematical rigor. Such violations instead raise what can sometimes be difficult questions for those interested in a rigorous reconstruction of the arguments of physics.

The remarkable thing, however, is that in spite of the difficult questions that such violations raise, mathematically unrigorous arguments can – at least, when things go well – settle questions of physics with great decisiveness, obliterating opposing schools of thought. It is strange that the standards of rigor and exactness lying at the heart of the mathematical method should be so dispensable in this way when we try to apply mathematics to the physical world. Some good examples of this are provided by Steiner in [90], where he invites philosophers to grapple with the broader epistemological puzzles that this phenomenon raises. I wish to accept Steiner's challenge, and assess what our attitude to all of this should be.

There are at least three views about what role the mathematicians' standards of rigor ought to play in physics. According to the first view, the physicist behaves irresponsibly when he formulates and allows himself to be moved by mathematically unpersuasive arguments, and should be criticized accordingly. According to this view, the physicist ought only to be moved by arguments that are mathematically rigorous. I will call this the *conservative* view of the role of mathematical rigor in physics.

According to the second view, until an area of physics reaches full maturity, physicists working in that area will likely be forced to traffic in arguments that are not fully mathematically rigorous. Given this fact, it would be unfair to criticize a physicist working in an immature area of physics for being occasionally moved by mathematically unrigorous arguments. However, once an area of physics reaches full maturity, we must demand that the physicist working in it *only* be moved by mathematically persuasive arguments. I will call this the *moderate* view of the role of mathematical rigor in physics; the difference between it and the conservative view is that it allows for non-rigorous argumentation in immature physical theories.

As always, variations on these views are possible. For instance, even the conservative view might allow for 'innocent' violations of mathematical rigor, where a violation is deemed innocent if it is the sort of thing which it is reasonable to conjecture *could* be made rigorous in a straightforward way. As mentioned already, care-free interchanges of limit signs are generally regarded (rightly or wrongly) as innocent in this sense. Likewise, the moderate view could allow for such violations in the case of a mature theory. Such minor modifications of the conservative and moderate view may be adopted at the reader's discretion, and do not affect the subsequent argument.

In addition to the conservative and moderate views, there is a third view about the role of mathematical rigor in physics – I will call it the *liberal* view – according to which it is inappropriate to demand that even the most mature areas of physics traffic exclusively in mathematically rigorous concepts and arguments. According to this view, even in the most mature areas of physics, it is still sometimes appropriate to be persuaded by an argument which is neither rigorous nor obviously rigorizable.[2]

Much mainstream philosophy of science rejects this liberal view outright. The positivist tradition, for instance, thought of physics as a discipline in which all theoretical concepts could be (and indeed, *had* to be) rigorously defined, and in which theories were given by axiomatic systems in which all facts of interest to the physicist could be deduced as 'theorems' in a logically rigorous manner. An example of this is Carnap's vision of science in the 'Aufbau' [16], in which he called for science to provide a logically rigorous 'construction' of its concepts from basic concepts. This is incompatible with the liberal view, which allows for the employment of concepts which do not meet – and perhaps cannot meet – the mathematician's standards of well-definedness.[3] Much subsequent philosophy of science also finds itself

[2]This, of course, is not to suggest that a rigorous physics is *impossible*, or never desirable.

[3]Note, however, that Carnap was probably a moderate rather than a conservative on the role of standards of meaningfulness in the sciences. See pp288-289 of [14]: *'Thus, the formation of the constructional system is the first aim of science.* It is the first aim, not in a temporal, but in a logical sense. The historical development of science does not have to postpone the investigation of an object until this object is placed within a constructional system. For objects on higher levels, especially for biological and cultural objects, science must not wait for this to take place, if it does not want to forgo, for a long time, the development of these essential fields with their important practical applications.'

clinging to the conservative – or at best moderate – view of the role of mathematical rigor in physics. On p. 111 of [43], for instance, Hempel tells us that he views scientific theories, in their 'advanced stages', as being given by 'deductively developed axiomatized systems'. In his influential essay 'Studies in the Logic of Explanation' [43], Hempel also argues that in any scientific explanation, the explanandum must be *logically deducible* from the explanans in order to be genuinely explanatory.[4] This goes against the liberal view, which allows for the possibility of an explanation which cannot be made cogent as a piece of logic, insofar as it involves concepts which are not rigorously defined, and not obviously rigorously definable.[5] Furthermore, even Popper argued that all physics should aim towards a 'rigorous system' – see in particular section 16 of [72], where Popper argues that this is important in order that we be able to decide unambiguously when a theory has been falsified by an observation.

Yet despite all this bad press, the liberal view enjoys great support amongst physicists, who often make disparaging remarks about the insistence on mathematical rigor, even in the most mature areas of physics. In their classic textbook [54], Landau and Lifshitz declare that *'no attempt has been made at mathematical rigor in the treatment, since this is anyhow illusory in theoretical physics . . . '*. In [3], Mandelbrot, replying to a comment by P. Anderson, says that *'. . . Anderson describes mathematical rigor as 'irrelevant and impossible'. I would soften the blow by calling it 'besides the point and usually distracting, even where possible.' '* (p 17.) In [12], Bridgman declares *'. . . the commitment of the physicist to the use of mathematics itself constitutes, paradoxically, a renunciation of the possibility of rigor.'* (p 227.) One also finds many well-read physicists citing Heaviside [42], who in reply to being criticized for the employment of unrigorous mathematical methods, rhetorically asked *'shall I refuse my dinner because I do not fully understand the process of digestion?'* (§225, p 9.) Finally, although the attitudes towards mathematics of physicists such as Feynman and Bohr are complicated enough to deserve a discussion on their own, it is clear that their views,

[4]See p247 of Hempel's [43]: 'The explanandum must be a logical consequence of the explanans; in other words, the explanandum must be logically deducible from the information contained in the explanans; for otherwise the explanans would not constitute adequate grounds for the explanandum.'

[5]In the main examples of nonrigorous mathematics I will be considering in this paper, one *can* conceive the relevant arguments and concepts as logically rigorous, so long as one is prepared to admit either inconsistent sets of premises or self-contradictory concepts. But given Hempel's insistence that consistency be a basic requirement of the axioms and concepts of a scientific theory, the tension between Hempel and the liberal view remains.

although different from each other, were species of the liberal view.[6] [7]

In this chapter, I wish to side with the physicists – and against Carnap et al. – by developing and defending the liberal view of the role of mathematical rigor in physics, according to which mathematically unrigorous arguments can sometimes be perfectly persuasive as mature pieces of physics. Specifically, I shall claim that arguments involving non-rigorous mathematical concepts (i) can be both natural and useful, so long as one respects some corresponding set of inferential restrictions, and (ii) can still justify their conclusions. These two claims provide further support (I shall argue) for the Principle of Inferential Restrictiveness, and the Deduction Failure Thesis.

Before going on, it must be emphasized that the liberal view does not assert that *any* violation of mathematical rigor ought to be permitted in physics, but only that *some* such transgressions ought to be. In order to argue for the liberal view, it therefore suffices to isolate just one important type of violation of mathematical rigor found in the physics texts, and defend its use. That will be the main job of the following sections.

In the next section, I will identify the type of violation that the present paper is to focus on – broadly speaking, the employment of mathematically ill-defined concepts within the context of an inferentially restrictive methodology. (This is intended to include the employment of mathematically self-contradictory concepts.) I will show how the physicist's use of delta functions and path integrals may be interpreted as violations of rigor of this sort that nevertheless contribute to physically persuasive arguments. This will help to demonstrate that the liberal view is indeed implicit in the methodology of physics as it has historically been practiced, thereby providing us with evidence for the Principle of Inferential Restrictiveness.[8] Merely showing that the liberal view is implicit in the physicists' practices, however, still leaves open the *normative* question of whether the physicist is *entitled* to behave in this way. In the fourth section, I will argue that the practice of physics is not misguided in this

[6] On pp 93-94 of [32], for instance, Feynman distinguishes mathematical rigor from physical rigor, asserting that they do not necessarily coincide. This attitude towards mathematical rigor can also be found throughout his famous undergraduate texts [33].

[7] For a discussion of Bohr's aversion to mathematics, see pp259-262 of Beller [9].

[8] Of course, this is not to say that all violations of mathematical rigor found in physics are of this sort, or that any violation not of this sort cannot be defended. To the contrary, I am quite convinced that there are radically different sorts of transgressions of mathematical rigor whose use in physics can be defended – but not necessarily by the arguments that will be given in this paper. In this paper I have simply tried to isolate one important class of violation, and defend it in a way that usefully illuminates some interesting problems about the relation between mathematics and the physical world.

way, and that the physicist is indeed entitled to pursue a methodology in which he permits himself to be moved by arguments violating the usual standards of mathematical rigor in the specific way described. In the fifth section, I will then dispel some general worries one might have about the employment of mathematically unrigorous methods in physics. In doing so, I aim to show that the demands placed on the physicist by the conservative and moderate views defined above are not reasonable.

3.3 REJECTING INFERENTIAL PERMISSIVENESS.

In this section, I wish to identify and discuss a specific class of unrigorous argument that one generally finds tolerated in the physics literature. The type of reasoning I am interested in involves concepts with ambiguous, or even self-contradictory properties. Specifically, I am interested in cases in which such concepts are employed by physicists with the proviso that certain otherwise mathematically acceptable forms of argument involving these concepts be prohibited. Prohibitions of this sort – which are, of course, nothing other than inferential restrictions – act as a way to control any damage that might arise from the ill-definedness or inconsistency of the concept in question. I will focus on two examples of this - the physicists' use of the delta function, and the use of path integral measures in quantum mechanics.

One might think that if physics involves inferential restrictions, then it must be all the *more* rigorous for it – after all, when we think of nonrigorous reasoning, we generally think of reasoning which goes *beyond* the argument forms generally accepted by the mathematician. Surely being more restrictive than the mathematician can only improve things.

But not all unrigorous reasoning need involve such things as explicit violations of the rules of the predicate calculus. One can sometimes engage in reasoning with a self-contradictory concept or inconsistent set of premises by deliberately abstaining from rules or patterns of inference which would obviously lead from the concept or premises in question to a contradiction. Insofar as such reasoning involves concepts or premises which *could* lead to contradictions in fairly obvious ways were certain rules of inference to be exploited, the mathematician would not describe such reasoning as rigorous, regardless of whether the reasoner explicitly abstains from such rules of inference or not. In such cases, one reasons unrigorously not so much by violating the laws of the predicate calculus, but rather by refusing to take such laws seriously enough – specifically, one reasons unrigorously insofar as one employs concepts and premises in a way which overtly flies in the face of such things as no-go or impossibility theorems. The earliest attempts at exploiting infinitesimals, for example, were criticized as unrigorous precisely on these grounds, and generally not on the grounds that such things as the law of non-contradiction had been violated. The sorts of unrigorous arguments in physics I wish to consider are precisely of this sort.

Note again that I need not claim that *all* instances of nonrigorous argument in physics are of this form. In order to justify the liberal viewpoint, it suffices to show that there are *some* mathematically unpersuasive arguments which ought to be regarded as persuasive pieces of physics. Different types of nonrigorous arguments in physics may well require different justifications; and there are perhaps arguments in physics whose nonrigorous nature cannot be justified at all. But so long as the rejection of inferential permissiveness is regarded as a legitimate methodological tool in physics – even in the case of mature theories – there will be a type of nonrigorous argument in physics that *can* be justified. This is enough to prove the liberal view correct.

Let us now turn to my primary examples of the phenomenon in question – the physicists' use of delta functions, and the use of path integral measures in quantum mechanics.

In many expositions of quantum mechanics, the delta function $\delta(x)$ is used to define wavefunctions of localized states in the position representation. It satisfies the defining relations:

$$\int_{-\infty}^{+\infty} \delta(x)dx = 1 \quad \text{and} \quad \int_{-\infty}^{+\infty} f(x)\delta(x)dx = f(0)$$

for any L^2 function f.[9] It is easily shown, however, that no such function $\delta(x)$ exists. For let f and g be L^2 functions such that $f(x) = g(x)$ almost everywhere, but $f(0) \neq g(0)$. Then, since $f(x)\delta(x) = g(x)\delta(x)$ almost everywhere,

$$\int_{-\infty}^{+\infty} f(x)\delta(x)dx = \int_{-\infty}^{+\infty} g(x)\delta(x)dx.$$

[9] A complex-valued, measurable function f is L^2 just in case $\int_{-\infty}^{+\infty} |f(x)|^2 dx < \infty$.

But the left hand side of this equation is just $f(0)$, and the right hand side $g(0)$, contradicting $f(0) \neq g(0)$.[10]

Delta functions are involved in many calculations in quantum mechanics. Given that the delta functions are, at best, a fiction, one might then wonder why we should trust the calculations of quantum mechanics involving them. Happily, there is a straightforward answer – so long as one keeps all delta functions inside integral signs, one cannot generate a contradiction.[11] Fortunately, this solution applies to quantum mechanics. In calculating, for instance, the expectation value of an observable for a state localized at a given point, all relevant delta functions *do* occur within integral signs, and the result obtained is not mere nonsense. As long as the physicist never asks questions about delta functions occurring outside integral signs – for example, as long as he never asks for the value of the delta function at a given point – his calculations will not be affected by mathematical contradictions such as the one identified in the previous paragraph.

It might be suggested, however, that in contrast to what I have said, the *real* reason that delta functions do not lead us into disaster is *not* that physicists knowingly refrain from interrogating delta functions outside integral signs, but rather that delta functions may be treated as a more abstract class of mathematical entity known as *distributions*.[12] Once we

[10] One might try to avoid this contradiction, by insisting that the delta function satisfy the defining relation $\int f(x)\delta(x)dx = f(0)$ only for $f \in L^2 \cap C^\infty$, where $f \in C^\infty$ iff f is infinitely differentiable. However, this does no better, as the following argument shows:
Fix real numbers $a < b$ such that $0 \notin [a, b]$, and assume that for some $r > 0$, $|\{x \in [a, b] : \delta(x) \geq r\}| = s > 0$. Select $[a', b']$ such that $[a', b'] \subseteq [a, b]$, $a \neq a'$, $b \neq b'$, and $(a' - a) + (b - b') < s$. Choose an f such that

1. $f \in L^2 \cap C^\infty$,
2. $f(x) \geq 0$ everywhere,
3. $x \notin [a, b] \to f(x) = 0$
4. $x \in [a', b'] \to f(x) \geq 1$.

This is always possible. Now $|[a', b'] \cap \{x \in [a, b] : \delta(x) \geq r\}| > 0$, so using properties 2., 3., and 4., $\int f(x)\delta(x)dx > 0$. But $f(0) = 0$, and so using 1., we have a contradiction with the defining relations for a delta function.
From this, we can conclude that for all $[a, b]$ not containing 0, and for all $r > 0$, $|\{x \in [a, b] : \delta(x) \geq r\}| = 0$. It follows that for all $[a, b]$ not containing 0, $\delta(x) = 0$ almost everywhere on $[a, b]$. From this we may conclude that $\delta(x) = 0$ almost everywhere on $\mathbb{R}$, from which it follows that for all $f \in L^2 \cap C^\infty$, $f(0) = \int f(x)\delta(x)dx = 0$, which is a contradiction.

[11] I will not provide the details of the argument for this here, but note that the arguments just given for the non-existence of the delta function revolve around considering the properties of the delta function outside the context of an integral.

[12] To be fully rigorous about the use of delta functions in quantum mechanics, one would also have to invoke such things as spectral theory, as demonstrated by von Neumann in [95]. The full details of this are not important for what follows, however.

abandon the idea that the delta function is actually a function and embrace the idea that it is a distribution, the threat of inconsistency vanishes. This, it might be suggested, is what saves the calculations of quantum mechanics from contradiction.

If one thought, however, that this was the real reason that delta functions do not lead physicists into disaster, one would have to be surprised – and perhaps even a little perplexed – by the fact that most quantum mechanics texts use delta functions extravagantly, without explaining in any satisfactory level of detail the fact that they are to be properly understood as distributions rather than functions. If it is this critical fact about delta functions that saves almost every quantum-mechanical calculation from meaninglessness, one would expect the correct interpretation of delta functions to be a matter of great urgency in the physics texts. But no such sense of urgency is to be found. What one finds, instead, is a curious sense of indifference. Because of this, I do not think that the existence of distribution theory on its own can fully account for the physicist's nonchalant attitude towards delta functions.

In fact, a closer look at the physics literature shows that the rigorization of delta functions by Schwartz in his 1945 paper [84][13] was an event of little significance for physicists, who generally felt long before then that the situation with delta functions was already well under control.[14] For instance, Dirac's textbook [26] – one of the classic texts of early quantum mechanics, first published in 1930 – showed very little anxiety over the use of the delta function, in spite of the fact that Dirac took the unsoundness of the delta function as a mathematical entity to be perfectly obvious to everyone. In this work, Dirac's only real comment indicating that there was something unusual about the delta function occurred in §15, where he wrote:

> *'... $\delta(x)$ is not a function of x according to the usual mathematical definition of a function, which requires a function to have a definite value for each point in its domain, but is something more general, which we may call an 'improper function' .... Thus $\delta(x)$ is not a quantity which can generally be used in mathematical analysis like an ordinary function,* ***but its use must be confined to certain types of expression for which it is obvious that no inconsistency can arise.'*** (Emphasis mine.)

These few remarks suggest that once one commits to avoiding the delta function in mani-

[13] For a little more detail on the history of the rigorization of delta functions, see pp 313-314 of Jammer's [47].

[14] It should be noted that the mathematical physicist Von Neumann was an exception to this. Von Neumann felt the introduction of delta functions introduced 'great mathematical difficulties' ([95], page 31) that he took very seriously. This, however, was a minority view.

festly troublesome contexts – a commitment which Dirac thinks can be honored in quantum mechanics – no questions about the delta function worthy of the physicist's attention remain. For Dirac, then, it is going inferentially restrictive that resolves any serious problems surrounding the use of the delta function. Furthermore, this view seems to be quite typical of the physics literature of the time. In light of this, it should be no surprise that the work of Schwartz had little effect on physicists, almost all of whom – much like Dirac – felt that there was really no great problem about the delta function to begin with, once the commitment to an inferentially restrictive methodology was made.

In no way did physicists change their tune after Schwartz's work of 1945. Even after the advent of distribution theory, most physicists were still happy to stick with the thought that it is judicious use of the delta function, rather than the existence of distribution theory, that saves quantum mechanics from disaster. On p67 of Dick and Wittke's 1960 textbook [24], we are told, for instance:

> *'The Dirac delta function is meaningful only under integral signs, where the limiting technique can be used.'*

Comments like this continue to pervade physics texts even to this day, in spite of the fact that a rigorous treatment of the delta function is now available. One must be struck by how little the lessons of distribution theory have impacted the physics community. Even Messiah's textbook – which actually contains a brief appendix on distribution theory – did nothing to challenge the view that the unrigorous, inferentially restrictive approach to delta functions was perfectly adequate. In Chapter V §8, for instance, Messiah tells us:

> *'. . . . In fact, no such function of p' and p'' possesses the desired property. Nevertheless, if one does not bother too much about mathematical rigor, one can, following Dirac, make use of the "singular function" $\delta(x)$ defined by the property . . . '*

where one of the usual definitions follows. Moreover, Messiah's appendix on distribution theory – one of the only in the quantum mechanics literature – is so cursory that one can hardly imagine it being useful to anyone seriously interested in the project of a rigorous reconstruction of quantum mechanics free from the delta function. So, in spite of Messiah's explicit acknowledgment of distribution theory, one cannot extract from his book the lesson that distribution theory is something that must be studied by any quantum theorist with a conscience who wants to sleep soundly at night.

In fact, the only thing that seems to separate the more careful from the less careful quantum mechanics texts is *not* the extent to which they discuss the details of the theory of distributions, but rather the extent to which they at least make some effort to inculcate the right set of *commitments* about the use of delta functions. No fundamental change occurs when one looks at papers published in the physics journals, instead of texts – the attitude to the delta function there remains the same. So even though after 1945 physicists were *able* to take an inferentially permissive approach to the delta function by invoking distribution theory, they nevertheless generally remained quite content to stick with their naive, unrigorous conception of the delta function, and pay the price of inferential restrictiveness. This attitude, which had been established well before 1945, was essentially unaffected by the discoveries of Schwartz.

All of the above suggests that in the physicist's view, the rigorization of the delta function was simply irrelevant to the question of the soundness of the usual calculative techniques of quantum mechanics. (Of course, the claim might be made that physicists were *wrong* in caring so little about the foundational issues surrounding the delta function. We shall turn to justificatory issues in the next section.) The use of the delta function in quantum mechanics shows that physicists are sometimes perfectly happy to knowingly exploit self-contradictory concepts – *even when rigorous alternatives are known to be available.* The physicist knows, of course, that this compromises his ability to adopt an inferentially permissive methodology. But the lack of interest in questions about the rigorization of the delta function seems to indicate that the loss of inferential permissiveness – at least in this case – is not regarded as any big loss. Physicists have been quite comfortable leaving their theories in a form in which one cannot treat them inferentially permissively. Insofar as physicists are comfortable reasoning inferentially restrictively in such cases, we have evidence for the Principle of Inferential Restrictiveness, and insofar as physicists are happy to leave their theories in a form in which inferential permissiveness is not an option, we have evidence for the Deduction Failure Principle.

Our next example of unrigorous reasoning in physics is the use of the path integral, which was invented and used by Feynman in order to explain a new way of thinking about the basic concepts of quantum mechanics. If a and b are space-time points, and $K(a, b)$ is the amplitude for a particle to go from b to a in the quantum system whose classical analogue has Lagrangian $\mathcal{L}$, then in the path integral formalism the identity

$$K(b,a) = \int_a^b e^{\frac{i}{\hbar}\int_{t_a}^{t_b}\mathcal{L}dt}\mathcal{D}x(t)$$

holds, where $\mathcal{D}x(t)$ is the 'path integral measure' over the set of possible paths from a to b through classical phase space. The path integral method gives an entirely different way of doing quantum-mechanical calculations from the usual operatorial methods. It may be given an intuitive motivation that the operatorial approach to quantum mechanics lacks. It also has the added feature of making certain numerical calculations in quantum mechanics more tractable.

It is unclear, however, how to construct the underlying measure $\mathcal{D}x(t)$ relative to which the path integral is to be defined. In the case of imaginary time (i.e., Euclidean space), the path integral measure can be defined with the use of the Wiener measure, a measure used originally in the study of Brownian motion. (See Glimm and Jaffe's [39] for details of this.) Of course, in field theory, there are many situations in which useful calculations *do* involve imaginary time, but for a completely general treatment of path integrals, the Wiener measure is insufficient. In complicated situations, such as when one uses the path integral to describe interacting fields in Minkowski space (rather than Euclidean space), there is little guarantee that there really is a well-defined underlying measure that can be used in the path integral. In fact, there are even theorems showing that in certain cases one *cannot* define a path-integral measure satisfying the requirements of the physicist.[15]

Here, we cannot simply gesture to some mathematical theory the way we can gesture to the theory of distributions when questioned about delta functions. The question of how the path integral is to be understood in full generality remains open. Given this, one might expect to see the physicist expending great energy trying to clarify the precise mathematical meaning of the path integral. Curiously, we again find that this is not the case. The typical physicist shows little interest in the rigorous work mathematicians have done on the path integral. This situation bears a striking resemblance to the pre-1945 attitude towards the

[15]Cameron [15] shows that there are cases in which no measure exists that can be used to define the Feynman integral. For a discussion of this and other difficulties involved with the path integral, see Chapter 6 of Rivers [77]. Also, see the Appendix of this dissertation for a discussion and strengthening of these results.

delta function – physicists simply see no point in losing sleep over a problem that they largely take to be well under control.

But what makes the physicists think that the situation with path integrals *is* largely well under control? The answer is that as long as the uses to which the path integral are put are of a particular sort, problems of inconsistency can be avoided. The pre-1945 use of the delta function was generally regarded as unproblematic so long as the delta functions were only used in a particular context – i.e., in expressions to be used as integrands. Likewise, there are ways of limiting one's use of the path integral in such a way that the threat of inconsistency disappears.

One can, for instance, think of the path integral formalism as merely an efficient algorithm for the generation of perturbation theory in either quantum mechanics or field theory. If we treat the path integral formalism in this way, no threat of inconsistency arises (so long as we take for granted the consistency of perturbation theory itself). Furthermore, this treatment of the path integral formalism does not demand that we take the path integral measure $\mathcal{D}x(t)$ seriously, just as employing the delta function within an integrand does not presuppose that we take the delta function seriously as a function.

In addition to this, provided that the path integral satisfies the usual properties of an integral, along with the 'convolution' property

$$K(b,a) = \int_{x_c} K(b,c)K(c,a)dx_c,$$

one can actually re-derive the Schrödinger equation.[16] These properties of the path integral on their own are certainly consistent, and – because they may be used to derive the Schrödinger equation – are sufficient for doing the non-relativistic quantum mechanics of particles. This provides us with yet another way of demarcating the argument forms involving path integrals that may be used without fear of inconsistency, in spite of the meaninglessness of the path integral measure $\mathcal{D}x(t)$.

There is, however, an important point of difference between the delta function and path integral. Unlike the pre-1945 use of the delta function, it is unclear *exactly* how literally one can take the idea of a path integral without getting into trouble. We know that the

[16]See §4.1 of Feynman and Hibbs [32].

path integral may be used unproblematically to generate perturbation theory, and we know that the path integral cannot be taken completely literally as an integral supported by a well-defined underlying measure. There are, however, many 'degrees of seriousness' between these extremes with which one may want to take the idea of the path integral, and for which it is generally *not* clear whether one is on safe ground or not. If the physicist wants to go beyond the safest ways of treating the path integral, he must make conjectures about consistency which he may well be forced to rescind at some later point. These conjectures are necessary because the conceptual and mathematical problems behind the path integral are just not as transparent as those with the delta function were prior to 1945. Thus, in the case of the path integral, the precise bounds of the physicist's inferential restrictiveness are not entirely clear. Yet in spite of this, the basic phenomenon in which I am interested remains unchanged – the physicist adopts a *mathematically incoherent* concept, and in order to avoid inconsistency, accepts certain inferential restrictions. The claim that the inferential restrictions in question successfully manage to avoid inconsistency may be a matter of conjecture rather than certainty, but the methodological fact that one permits mathematically incoherent concepts into one's discipline by going inferentially restrictive remains unchanged.

In summary, my central claim about the use of the path integral in physics is much the same as my claim about the use of the delta function in physics – it is a type of unrigorous argumentation which does not bother the physicist, insofar as he is happy to adopt an inferentially restrictive methodology that protects him from unpleasant consequences.[17] The willingness to go inferentially restrictive in turn explains the indifference one finds amongst practicing physicists to the progress the mathematician has made or continues to make in rigorizing the concept of the delta function or path integral.[18] Insofar as physicists are happy to be moved by arguments which are neither rigorous nor obviously rigorizable – even in a highly mature discipline such as non-relativistic quantum mechanics – it is clear that the

[17] In [67], Norton has suggested that something very much like adopting an inferentially restrictive methodology may be useful for sidestepping certain inconsistencies in Newtonian Gravitation Theory. For a careful discussion of the inconsistency in question, see also Malament's reply [57] to an earlier version of Norton's paper.

[18] Of course, in the sense that the mathematics literature might help to validate or refute the physicist's conjectures about the consistency of certain argument forms involving path integrals, the physicist might have something to gain from the mathematician. This, however, is generally not the direction of mathematical research into the path integral.

liberal view outlined in the previous section does indeed permeate the practice of physics.[19] Thus, the attitude of the physicist towards delta functions and path integrals provides us with support for the Principle of Inferential Restrictiveness, according to which physicists do not feel bound to accept all the mathematical consequences of assertions that they accept. Furthermore, insofar as physicists are happy to leave bodies of knowledge in a form in which they are not closed under arbitrary mathematical consequences, we also have evidence for the Deduction Failure Principle.

In this section, I have been content to make descriptive claims about the practice of physics. A natural response to all this, however, is that in spite of actual practice, physicists should *not* be persuaded by unrigorous arguments – especially in mature theories – and likewise should *not* be so nonchalant about adopting an inferentially restrictive methodology. In order to defend the liberal view against these attacks, we must turn our attention to justificatory matters such as these.

3.4 THE PROBLEM OF JUSTIFICATION.

In analogy with the previous chapter, we consider now whether one can ever be *justified* in asserting the conclusion of a mathematically unrigorous argument. To put the matter more pointedly: *why* should we trust arguments that invoke delta functions (for instance), given that we can use the contradictory properties of delta functions to deduce *anything*? If we cannot answer such a question, then my suggestion that physicists are comfortable with nonrigorous concepts proves little.

Let us consider a first answer: in an inferentially restrictive methodology, each step of every argument is still logically valid. What distinguishes inferentially restrictive methodologies is the commitment to *avoid* certain otherwise legal inferences – but that does not mean

[19] It is important to note that the inferentially restrictive methodology of physics does not make it impossible for the blind pursuit of mathematics to lead to physically significant discoveries. A good example of progress in physics being made in this way would be the discovery of anti-particles associated with negative energy solutions of the Dirac equation. Other examples of discoveries in physics arising from purely formal or mathematical considerations are discussed by Steiner in [91]. Nothing about the adoption of an inferentially restrictive methodology prevents novel applications of purely mathematical ideas in physics; even though their creators will perhaps be required to defend such ideas with greater vigor than they would were the methodology of physics inferentially permissive.

that the inferences one actually performs are illegal. There is no question, then, that the conclusion of a calculation or argument is entailed by the premises in the sort of nonrigorous argument I have been focusing on.

Nevertheless, an opponent might object that we are never justified in ignoring an otherwise valid logical inference. So even though a calculation might lead to a true conclusion, the fact that if certain barred inferences were to be permitted, a contrary result would be obtained, is enough to make it the case that the calculation in question does not really justify the answer. One can, after all, prove any conclusion one wants by starting with an inconsistent set of premises and cleverly restricting inferences.

Of course, those who work in paraconsistent logic would argue to the contrary that one *can* justify a conclusion from an inconsistent set of premises. In many obvious ways, reasoning in an inferentially restrictive manner is like paraconsistent reasoning.[20] The existence of approaches to logic in which a conclusion may be justified on the basis of inconsistent premises should therefore give my opponent pause.[21]

But I can also give another answer, similar to that given in the previous chapter. It will help to consider a slight variant of the three light system introduced there. Recall that the three-light system consists of three lights A, B, and C, each of which flashes on, very briefly, at various times. Let us suppose that at time $t = 0$, all the lights flash on, light B flashes on every .002 seconds afterwards, light C flashes on every .003 seconds afterwards, and light A flashes on at times $t = 0, 1.005, 2, 3.003, 4, 5.001, 6, 7.005, 8, 9.003, 10, 11.001, 12....$ Call these set of laws L. Only the times at which A flashes differ from the example of the last section, and the difference is small. The times of A's flashings are chosen such that at any time $t > 0$, whenever A flashes, exactly one of B or C flashes with it. Furthermore, for $t > 0$, the first time A flashes, C also flashes, the next time A flashes, B flashes, the next time A flashes, C flashes, and so on. Which light flashes along with A alternates from B to C.

A scientist who does not know all these details, makes the following observations:

(O1) Each individual light has a period τ such that the light seems to flash on roughly every

[20] An important difference is that in the case of paraconsistent logic the restrictions on inference rules are usually based on the *logical form* of those rules – in the case of nonrigorous reasoning in physics, the restriction depends instead on the *content* of the propositions involved.

[21] For a further discussion of reasoning with inconsistent scientific theories, see Smith [88] and Brown [13].

τ seconds, where $\tau_A \sim 1, \tau_B \sim 0.002$, and $\tau_C \sim 0.003$.

(O2) At $t = 0$, all three lights flash on.

(O3) Whenever light A flashes on, exactly one of light B or C also flashes on.

(O4) For $t > 0$, the first time A flashes, C also flashes.

(O5) For $t > 0$, if A flashes along with B, then the next time A flashes, C flashes with it.

(O6) For $t > 0$, if A flashes along with C, then the next time A flashes, B flashes with it.

Note that if we assume that each light has an *exact* period τ, then **(O2)-(O6)** cannot all be true, for essentially the same reasons given in the previous chapter.

What can the physicist do in such a situation, if, for some reason, he is unwilling to let go of the hypothesis that each light is exactly periodic? Let us consider the following approach.

The physicist recognizes that $\tau_A \sim 1, \tau_B \sim 0.002, \tau_C \sim 0.003$, that $2\tau_A$ is an integral multiple of τ_B, but that τ_A is not an integral multiple of τ_B. These last two requirements are easy enough to satisfy; they simply imply that τ_A/τ_B is a half-integer (i.e., has the form $n + 1/2$, where n is an integer.)

The physicist also recognizes that he must have that (i) $\tau_A, 3\tau_A, 5\tau_A, ...$ are all integral multiples of τ_C, while (ii) none of $2\tau_A, 4\tau_A, 6\tau_A, ...$ are integral multiples of τ_C. Of course, these requirements are incompatible. Nevertheless, the physicist can do quite well by simply saying that $\eta = \tau_A/\tau_C$ is a 'pseudo-integer', where a pseudo-integer η has the property that $\eta, 3\eta, 5\eta, ...$ are integers, while $2\eta, 4\eta, 6\eta, ...$ are not integers. Thus, the physicist summarizes his knowledge of the system as follows:

(T1) Each light is *exactly* periodic.

(T2) $\tau_A \sim 1, \tau_B \sim 0.002, \tau_C \sim 0.003$

(T3) Each light flashes on at $t = 0$.

(T4) Whenever light A flashes on, exactly one of light B or C also flashes on.

(T5) τ_A/τ_B is a half-integer.

(T6) τ_A/τ_C is a pseudo-integer.

Much like a delta function, the concept of a 'pseudo-integer' is entirely unrigorous; such things obviously do not exist.[22] Nevertheless, the physicist can reason quite successfully with

[22] If the reader thinks the concept of a 'pseudo-integer' is far-fetched, the reader should recall the concept of a Grassman variable as used to describe fermionic systems in the path integral formulation of quantum

‘pseudo-integers’, so long as he restricts the sorts of arguments in which they may be used. For instance, one may use the fact that τ_A/τ_C is a pseudo-integer to draw conclusions about whether C flashes at any given time that A is flashing, but nothing else.

As in the previous chapter, we make the following definitions: let T_1 be the set of claims that are consistent with L. For sentences $p, q \in T_1$, say that $p \leq_1 q$ iff q may be deduced from p and L, and define $p =_1 q$ iff $p \leq_1 q$ and $q \leq_1 p$. For any $p \in T_1$, let $[p]_1$ be the set of $q \in T_1$ such that $p =_1 q$. For $p, q \in T_1$, define $[p]_1 \leq_{[1]} [q]_1$ iff $p \leq_1 q$. Then $\leq_{[1]}$ defines a partial order on the set of equivalence classes $[p]_1$ with $p \in T_1$; call the resulting partially ordered set P_1.

Now let T_2 be the set of claims that (i) are consistent with **(T1)–(T6)**, and (ii) can be built up out of sentences of the form ‘lights A and/or B and/or C flash on at approximately time t’ (for positive real numbers t.) For $p, q \in T_2$, define $p \leq_2 q$ iff q is mathematically deducible from p, with all inferential restrictions being respected. Define $p =_2 q$ iff $p \leq_2 q$ and $q \leq_2 p$. Also define equivalence classes $[p]_2$ for $p \in T_2$ and an induced partial ordering $\leq_{[2]}$ on these equivalence classes as before, and call the resulting partially ordered set P_2.

As before, we have the following result:

Poset Homomorphism Result: $\forall p, q \in (T_1 \cap T_2)[([p]_2 \leq_{[2]} [q]_2) \rightarrow ([p]_1 \leq_{[1]} [q]_1)]$.

In this case, the Poset Homomorphism Result is a general result about the reliability of the ‘pseudo-integer’ concept (provided inferential restrictions are respected). As in the previous chapter, the Poset Homomorphism result may then be used to justify the conclusion of arguments involving the pseudo-integer concept.

In the case of delta functions and path integrals, we have similar results. In the case of the delta functions, the partial order T_1 corresponds to the set of claims consistent with the true laws of nature L, ordered by $p \leq_1 q$ iff q can be derived from p and L. The partial order T_2 consists of the claims consistent with quantum mechanics, ordered by $p \leq_2 q$ iff q can be derived from p and Schrödinger’s equation, subject to (i) the inferential restrictions pertinent to the use of the delta function/ path integral, and (ii) the inferential restrictions pertinent to the fact that quantum mechanics can only be expected to yield accurate results

mechanics. Grassman numbers are often taken to be non-zero real numbers which anti-commute – i.e., if g_1 and g_2 are Grassman numbers, $\{g_1, g_2\} = g_1 g_2 + g_2 g_1 = 0$.

in certain situations and regimes. Given such a definition, the above Poset Homomorphism Result appears to hold.

We can say a little more about the delta function case. Order the set T_2 of claims consistent with quantum mechanics as follows: $p \leq_2^* q$ iff q can be derived from p and Schrödinger's equation, without any use of the delta function at all, subject only to the inferential restrictions pertinent to quantum mechanics itself. Define equivalence classes $[p]_2^*$ and a partial order $\leq_{[2]}^*$ on these equivalence classes as before. Then we have the following:

Special Poset Homomorphism Result:

1. $\forall p, q \in (T_1 \cap T_2)[([p]_2^* \leq_{[2]}^* [q]_2^*) \rightarrow ([p]_1 \leq_{[1]} [q]_1)]$.

2. $\forall p, q \in T_2[([p]_2 \leq_{[2]} [q]_2) \rightarrow ([p]_2^* \leq_{[2]}^* [q]_2^*)]$.

Facts (1.) and (2.) obviously entail the Poset Homomorphism Result. Fact (1.) simply says that, in its intended domain, quantum mechanics (without delta functions) is indeed a successful theory. Fact (2.) says that implications in quantum mechanics involving the delta function may be captured by implications that do not involve the delta function. Fact (2.) therefore holds in virtue of the fact that the delta function *can* be made rigorous.

Thus, when one introduces a set of inferential restrictions associated with a concept that *can* be rigorized, the homomorphism in the Poset Homomorphism Result will be factorizable into two homomorphisms, in the way shown in the Special Homomorphism Result above. However, what justifies the use of delta functions is not the fact that such a factorization is possible, but rather the mere existence of the homomorphism introduced in the Poset Homomorphism Result in the first place. The rigorizability of a mathematical concept is not a factor in justifying its use – all that matters is that, when appropriate inferential restrictions are introduced, inferences involving the concept in question 'track the truth' in the way articulated in the Poset Homomorphism Result.

Thus, we have justified the use of non-rigorous mathematical arguments in physics in much the same way that we justified he use of idealizations (with associated inferential restrictions) in the previous chapter.

3.5 FURTHER COMMENTS ON JUSTIFICATION.

At this stage, an opponent might declare that the liberal view of the role of mathematical rigor in physics has still not been justified. According to the liberal view, it is sometimes appropriate to be persuaded by an argument which is not mathematically rigorous even in the case of mature theories. An opponent might declare, however, that any area of physics in which it is acceptable to be persuaded by a mathematically unrigorous argument is, of necessity, immature. Thus, we have at best given an argument for the moderate view; but certainly not the liberal view.

To further motivate this objection, recall that in [52], Kitcher argues that one of the things that makes rigor necessary in mathematics is that it helps us to understand why our mathematical calculations work. Here, Kitcher is thinking of the calculations that lie behind *mathematical* problem solving, such as finding roots of equations, calculating maxima and minima, etc., rather than problem solving in physics. But there is an obvious analogue of Kitcher's position in the world of physics problem solving, according to which rigor is necessary in physics precisely because it helps us to understand why our *physics* calculations work. If rigor is required in physics for this reason, then the presence of non-rigorous arguments indicates the absence of a certain type of understanding. This, in turn, indicates a type of immaturity. So any area of physics in which physicists are persuaded by non-rigorous arguments is necessarily immature.

What is meant when it is said that we do (or do not) understand why a calculation works? One can imagine a calculation involving strange steps, in which is it simply not clear how to view the argument as exhibiting the way in which the laws of nature, as they are ordinarily understood, lead to the result or phenomenon in question. (Perhaps some physical results obtained by the method of analytic continuation in complex analysis are of this form.) In such cases we confess that we do not understand why the calculation works. By contrast, there are calculations in which each step follows inevitably from the one before, and the calculation as a whole can be viewed as a natural expression of the way in which simple, well-understood, processes come together to produce the physical result in question. (Many simple calculations in classical mechanics, for instance, are of this form.) In cases such as these, we claim to understand why the calculation works.

To what extent, then, is mathematical rigor needed to produce this feeling of understanding? There is no denying that sometimes attention to rigor can help to produce understanding. There is also no denying that, when abused, nonrigorous mathematical methods can hinder understanding. But these are nothing more than rules of thumb. Rigorous methods can sometimes be entirely unilluminating, and unrigorous mathematical calculations can sometimes bring with them a deep understanding of how underlying processes work together to produce the phenomenon under investigation. (Path integrals in quantum mechanics are an excellent example of this, in which Schrödinger's equation is connected with the classical concept of 'action'.) In the case of nonrigorous reasoning, part of this understanding will involve appreciating that certain sorts of mathematical interrogation should not be regarded as physically significant, and should consequently be ignored. But it is difficult to see how this hinders understanding, as the case of path integral methods in quantum mechanics shows.

To demand that a mature physical theory use only rigorous concepts is therefore to place an inappropriate demand on what makes a theory 'mature'. To call a theory 'mature' is to comment on its depth, scope, and the way in which it enables us to view complex physical phenomena as springing forth from simpler physical facts. None of this demands mathematical rigor as a prerequisite, and I therefore think that this attack on the liberal view fails.

Let us consider another concern. One might worry that by allowing mathematically unrigorous arguments into physics, all sorts of absurdities become permissible. For instance, imagine that an experimental result is obtained in conflict with some theory. Let us further imagine that by deliberately not carrying the 1 in some particular addition problem involved in the theoretical calculation, agreement can be restored between theory and experiment. We might then imagine someone declaring that the experiment in question has shown that a new, unrigorous understanding of addition is needed in physics, and that the old theory is still unfalsified. As it stands, this is lunacy. But the worry is that if we allow *any* mathematically unrigorous arguments into physics, we will have no basis on which to condemn this sort of move.

The same worry can also be a expressed a little differently. It might be suggested that the main reason certain physicists were so nonchalant about the nonrigorous mathematical methods of quantum field theory (especially in the days when renormalization was not as deeply understood as it is today) is that the theory had enjoyed truly extraordinary empirical success, and that this is really all physicists ought to care about – coming up with theories that accurately tell us what we will see in our particle accelerators. Thus, the use of nonrigorous mathematical methods in physics is defended by pointing to the empirical success of relevant theories. The worry, however, is that if all that matters in physics is devising a calculus which can predict what we will see in our particle accelerators, then all sorts of ad-hoc physics is possible – such as the absurd example given in the previous paragraph.

My reply is that while some people might naively defend the use of unrigorous mathematics in physics by pointing to empirical successes and shrugging their shoulders, I have gone to great lengths to provide an *alternative* way of defending the use of unrigorous mathematical methods in physics. My defense certainly does not commit me to the view that all that matters to the physicist is that he come up with a calculus that can predict what will be seen in the particle accelerator. I fully agree that it is the job of physics to do more than this. The physicist does not only want to see his calculations work – he wants to understand *why* his calculations work. This requires, amongst other things, that he be able to view his calculations as articulating some set of physically intelligible processes that conspire to produce the empirical phenomenon in question. Our desire to *understand* the success of our calculations in this way is an important criterion of success for a theory in physics; and it is because the example given two paragraphs ago fails so drastically to meet this requirement that we correctly think of it as lunacy.

There is an obvious distinction to be drawn between an unrigorous physical theory for which we feel we understand why its calculations work (such as the path integral formulation of quantum mechanics), and an unrigorous physical theory for which we do not (such as the 'pseudo-addition' example given earlier). Insofar as it is possible to draw this distinction, the final worry of this section can be easily defused – the lunatic theories under consideration will not meet the criteria of a good physical theory for reasons ultimately unrelated to their mathematically unrigorous methodology.

The sorts of justificatory worries that I have presented in this and the previous section do not speak against the possibility of using mathematically incoherent concepts in persuasive physics arguments. In fact, such arguments can be found throughout physics. The fact that many physicists are undisturbed by this speaks in favor of the Principle of Inferential Restrictiveness. Additionally, the fact that physicists are committed to bodies of knowledge that are *not* closed under arbitrary mathematical inferences (insofar as arbitrary mathematical inferences can lead to contradictions) also provides us with important evidence in favor of the Deduction Failure Principle.

4.0 UNPHYSICAL ARGUMENTS.

4.1 INTRODUCTION

In this chapter, I will discuss one thing (amongst many) that a physicist can mean when he declares an argument to be 'unphysical'. In the process, I shall find additional support for both the Principle of Inferential Restrictiveness and the Deduction Failure Principle.

Every now and again, a physicist will declare an assertion, argument, or even an entire research programme to be 'unphysical', or based on 'unphysical ideas'. There are a great many things a physicist can mean when he makes such a judgment – we will examine a few of the possibilities. In some cases, such a judgment can be uncontroversial and unproblematic, but in other cases subtle philosophical questions are raised in trying to understand exactly what is meant.

Let us consider some easy cases first. Often, an equation which models a given system has multiple solutions. Which solution must we pick? Sometimes certain solutions will have pathological features, rendering them 'unphysical' in a fairly straightforward way. For instance, a solution to an equation which assigns negative or imaginary mass to a particle, or which involves a violation of causality or the second law of thermodynamics, can generally be declared 'unphysical' and ignored without further ado, especially if other non-pathological solutions exist. Examples of this sort will not be my main focus.

Other judgments of 'unphysicality' are harder to understand. I shall briefly outline a few of these more difficult cases here, some of which will be developed further in the course of the chapter.

Consider the differential equation

$$\frac{d^2x}{dt^2} + \lambda(\frac{dx}{dt} - 1) = \Phi(x)P_T(t).$$

Here, Φ is an arbitrary function, and P_T is a periodic 'pulse function' with period T, equal to 1 on a short subinterval of $[0, T]$, and 0 elsewhere on $[0, T]$. Such a differential equation describes a periodically forced, damped particle in one dimension. This equation displays a rich variety of chaotic behaviors for a large range of values of λ and T. The set of λ and T for which we get such chaotic behaviors consists largely of open intervals – i.e., it is not just a set of isolated points which can be somehow dismissed as 'anomolous'. (For mathematical details, see [96].)

In particular, for a large, open set of values of λ and T, the dynamical system has a 'strange attractor'. [1] Strange attractors are subsets of phase space with bizarre, 'fractal-like' topologies, of non-integral dimension, generally containing subsets isomorphic (in various senses) to the Cantor set. (For more precise details, see [97].) Such features are not peculiar to the differential equation described above – continuum mechanics and fluid mechanics are full of differential equations whose solutions have strange attractors.

It is tempting to think that the kind of fractal structures associated with strange attractors do not correspond to any *real* subset of physical phase space; instead, it is tempting to think of such structures as 'unphysical'. For instance, in a review of Smith's book [89], Callendar writes [14]:

> *'Chaos theory idealizes nature, to be sure. But at least at first glance, it idealizes nature the wrong way around. Frictionless planes and such examples leave out 'irrelevant' detail; chaotic models add 'irrelevant' detail. They model real-world systems with intricate geometries in state space that no real system's temporal evolution could possibly follow. Yet despite all this 'surplus structure' scientists appear to use these models to describe, predict and explain physical behaviour. How is this possible?'*

[1] An 'attractor' is a subset A of phase space with the following properties:

(1) A is invariant under the dynamics; i.e., if the system begins in A, it will remain in A.

(2) There is some open neighborhood U of A (the 'basin of attraction') such that all trajectories that start in U are 'attracted' to A – formally, if $u(t)$ is a trajectory with $u(0) \in U$, then $lim_{t\to\infty}D(A, u(t)) = 0$, where $D(A, u(t))$ is the distance between $u(t)$ and the set A (i.e., $D(A, u(t)) = \inf_{a\in A} d(a, u(t))$, where d is the usual metric.)

(3) No proper subset of A satisfies (1) and (2).

A strange attractor is an attractor with non-integral dimension (for a discussion of the various notions of dimension one can use here, see [31] or Chapter 2 of [89].)

In declaring the 'intricate geometries' of chaos to be 'surplus structure', Callender expresses suspicion about the physicality of fractal structures. To be sure, he agrees that such structures are useful for explaining a large class of natural phenomena. In some cases, for instance, one can say that it is because a strange attractor has some bizarre topological property rather than some other bizarre topological property, that the dynamical system in question exhibits one sort of macroscopic behaviour rather than another. Bizarre subsets of phase space can therefore occupy a crucial role in scientific explanations. But in spite of this, Callender is tempted to say that such structures are 'unphysical', and correspond to nothing in nature. This concern can also be found in Smith, who asks *'How can an infinitely intricate structure like this possibly play an essential part in a competent scientific account of some natural phenomenon?* This is a suspicion that one also finds echoed amongst many physicists. Explanations such as 'it is because phase space contains a strange attractor with bizarre topological property X that the dynamical system will exhibit macroscopic behaviour Y' have consequently struck many as entirely 'unphysical'. What exactly is meant by 'unphysicality' in this case, however, is nothing as simple as in the case of negative mass or a violation of causality.

Another sort of case concerns solutions to the three body problem. Consider three masses in otherwise empty space, that interact with each other according to the usual Newtonian inverse square gravitation law. A great deal of energy has been spent by mathematicians identifying stable[2], periodic orbits that can occur in this context; such investigations go back to Euler, Lagrange and Poincare. Modern mathematicians have continued to produce more and more surprising possibilities for periodic orbits, some involving even larger, finite numbers of particles. This area of mathematics has produced an amazing repertoire of truly baroque, periodic dances that a large number of particles can perform, consistent with the inverse square law. (See Montgomery's [62] for further details.)

In spite of the beauty of these results, this is an area of research populated almost entirely by mathematicians, and in which physicists take a very small interest. There are surely many reasons for this lack of interest. I will focus, however, on one particular reason that some physicists give – namely, that such research takes the inverse square law 'too seriously', and

[2]Various notions of stability are used throughout this literature – for instance, see page 473 of [62].

therefore produces results that we have no real reason to think could ever be realized in nature. In this sense, such periodic, stable orbits are declared to likely be 'unphysical'.

Other interesting examples of 'unphysicality' are not too hard to come by in physics. For a long time, many prominent physicists felt that black holes were nothing other than an unphysical artifact of the mathematics of general relativity, that came from taking Einstein's field equations too seriously in the neighborhood of a singularity. Anyone who tries to apply Schrödinger's equation beneath the Planck scale, or Newton's law of gravitation to truly massive bodies will likewise be accused of unphysical argumentation. But again, such accusations of 'unphysicality' are not as easy to understand as the case of negative masses or violations of causality.

There can be little doubt that in each of these cases, the term 'unphysical' is used in a slightly different sense. I have no aspiration to give a completely general account of what it means to call something 'unphysical'. One might even wonder whether the term 'unphysical' is nothing other than a general term of derision, with no real unified, underlying meaning. I think, however, that there *is* something that is *often* meant by calling an argument 'unphysical', that applies to an awfully large number of cases – perhaps larger than might be expected. Spelling out and developing this sense of the term will be the goal of much of this chapter. This will involve another appearance of the concept of inferential restrictiveness, and with it, more evidence for both the Principle of Inferential Restrictiveness and the Deduction Failure Thesis.

I should remark that the goal of this chapter is *not* to try and take sides, and (for instance) try to settle the question of whether fractals really are 'physical' or not (whatever we decide that means). It is unlikely that any light can be shed on such questions using the tools of philosophy alone. Nor do I claim that physicists are *always* right – or even *often* right – when they declare something to be 'unphysical'. The example of black holes illustrates this point nicely. All I wish to explain is what, in certain situations, a physicist might *mean* in judging something to be 'unphysical'. Deciding whether, in any specific case, such a judgment is correct is not the sort of debate I shall participate in here.

In sections 2–3, I would like to consider some proposals for what is going on when a physicist judges something to be 'unphysical' in certain types of situations. These proposals will be shown to be deficient in various ways. In section 4 I will re-introduce the concept of inferential restrictiveness, in order to present my final analysis of 'unphysicality' in section 5. In addition, I will also draw connections with the Principle of Inferential Restrictiveness and the Deduction Failure Principle in section 5. In section 6 I will present some concluding remarks for both the present chapter as well as Part 1 as a whole.

4.2 SOME PROPOSALS.

4.2.1 Simple Proposals.

What is going on when a physicist declares an argument or assertion to be unphysical?

One view that I alluded to in the previous section is that when a physicist declares an assertion or argument to be 'unphysical', all he is doing is expressing his disagreement with it, without giving a reason. According to this view, to declare that fractals are unphysical is just to disagree with the explanations provided by chaos theory, without reason – for if the physicist *had* any such reason, he could just as easily cite *it*, rather than making the vague accusation of 'unphysicality'.

But this sort of view should be a last resort. The fact that the term 'unphysical' can be used in an uninteresting or even vacuous way does not mean that it cannot also be used in an interesting, non-trivial way. Until we have exhausted all such alternatives, the view suggested here is premature. Let us then consider some alternatives.

Here is another proposal. Consider the case of the three body problem. In declaring certain types of orbits 'unphysical', perhaps the physicist just means that such configurations are unlikely to be found in nature. So for instance, although such orbits are stable, the required initial conditions are extremely delicate, and so we might conjecture that the probability of finding such orbits in nature is vanishingly small. In this sense, these strange orbits are 'unphysical'.

But I do not think this analysis is very satisfying. All sorts of physical phenomena are exceedingly rare. Superconductivity and Bose-Einstein condensation do not occur outside the most controlled of laboratory situations, yet that is not enough to render such phenomena ‘unphysical’. Moreover, even before experimental successes with these phenomena, they were *not* generally dismissed as unphysical by physicists. Thus, ‘unphysicality’ must mean more than mere ‘rarity’ or ‘unlikelihood’. This is not to deny that the term ‘unphysical’ may sometimes be used to mean ‘rare’ or ‘unlikely’ – but in such cases, I shall simply suggest that the term ‘rare’ or ‘unlikely’ ought to be substituted for ‘unphysical’. Such cases will not be of interest here.

4.2.2 Perturbing the Microtheory.

Let us consider the three body problem again, and the question of whether the strange orbits constructed by mathematical physicists are the sorts of things that could possibly be found in nature.

One might argue that the constructions of the mathematicians in this area are ‘unphysical’ because they are not stable enough under small changes in the underlying microtheory. Specifically, these constructions presuppose the validity of Newton’s law of gravitation at all scales – a presupposition which we know to be false. On its own, this is no crime, as classical arguments for the possibility of orbits along conic sections presuppose this too, and we do not question their physicality. However, one might conjecture that unlike orbits that lie on conic sections, the delicate nature of the strange orbits is such that we have no reason to expect that they remain genuine possibilities when we perturb the microtheory – as is necessary to account for more subtle effects. It is because of this (conjectured) ‘instability’ under perturbation of the underlying microtheory that one might take such strange orbits to be ‘unphysical’.

But I do not think that this is a particularly good analysis of ‘unphysicality’. Physicists routinely accept arguments for which it is quite unclear how their conclusions would be affected by perturbations in the underlying microtheory. Consider the derivation of the ideal gas law from the ‘billiard ball’ model of a dilute gas — small changes in the way in which the billiard balls interact (for instance, the introduction of tiny inelasticities) can produce macroscopic changes in the subsequent equation of state. Likewise, any subsequent discovery that all electrons are *not* utterly identical would call into question arguments that exploit the Pauli Exclusion Principle to deduce, for instance, the properties of electron gases.

One might reply to this by pointing out that the ideal gas equation of state, and the properties of electron gases are well known from an *experimental* point of view, and so these examples demonstrate nothing. But this is besides the point. Physics is full of both good and bad theoretical arguments that support known empirical facts. Why should we think that the current theoretical arguments for the ideal gas law are *good* arguments? The answer *cannot* be – because these arguments are resistant to perturbations in the underlying microtheory; because such a claim is untrue. Resistance to perturbations in the underlying microtheory therefore cannot be adequate grounds for dismissing an argument as 'unphysical', for if it is, physics is full of persuasive, yet unphysical arguments; and the accusation of 'unphysicality' is no longer really worth fussing about.

Of course, if an argument is *blatantly* unstable under the most trivial perturbations of the underlying microtheory, then a physicist may have good reasons for dismissing it as unphysical. But I think such cases are quite rare. I consequently do not think that the proposal of this subsection gets us very far in understanding what might make an argument 'unphysical' in the eyes of a physicist.

4.3 ENERGY, SPACE, AND TIME REGIMES.

Let us consider another proposal. Any given law (Newton's law of gravitation, Schrödinger's equation, etc.) will have a certain domain of validity; that is to say, regions of energy, mass, length, time etc. such that the law will be good inside those regions, but much less accurate (if even meaningful) outside those regions. When a law is used outside such a region – or worse, in a region in which it is *known* to be fail badly – the resulting argument may be said to be 'unphysical'. This is the proposal I will consider in this subsection.

According to this proposal, it is 'unphysical' to explain the macroscopic behavior exhibited by a system that follows the macroscopic equation

$$\frac{d^2x}{dt^2} + \lambda(\frac{dx}{dt} - 1) = \Phi(x)P_T(t)$$

by citing the fact that some strange attractor has some arcane topological property, because the differential equation here is a *macroscopic*, classical equation, and the relevant topological properties of the strange attractor are fundamentally *microscopic* in nature. Likewise, the strange orbits that solve the three body problem require the law of gravitation to be true at all scales, and so we have no reason to think that they are physical. Similarly, for a long time the best evidence we had for general relativity was based on the scattering of photons and (relatively) light objects by large masses (usually stars) at large distances – and so certain types of arguments for the existence of black holes based on the validity of Einstein's equations near a singularity were not obviously physical. Attempts to apply Schrödinger's equation beneath the Planck scale, or Newton's law of gravitation to galactic clusters, are unphysical for analogous reasons.

Let me make two notes about this proposal. First, it does not aim to cover *all* uses of the term 'unphysical' – rather, it is content with identifying a widespread (though not necessarily universal) use of the term. Second, I reiterate that there is nothing about this proposal that suggests that when physicists judge something to be 'unphysical', they are necessarily always *right* – or even *sometimes* right – about it. All this proposal seeks to do is explain, in certain cases, the *meaning* of the physicist's assertion.

I think this proposal is a lot closer to the truth, but is still not quite right. To see why, consider the following two phenomena:

(1.) Take the differential equation

$$\frac{d^2x}{dt^2} + \lambda(\frac{dx}{dt} - 1) = \Phi(x)P_T(t)$$

discussed already. This a macroscopic equation – it is intended to describe the macroscopic behavior of a particle, and makes no attempt to incorporate quantum effects, and so on.

Note, however, that in virtue of being a differential equation, it manages to say something about the macroscopic behavior of the particle only by stating a relation that holds between infinitesimal changes in x and infinitesimal changes in t. Now, even if the differential equation in question models the macroscopic behavior of the particle well, the relation between infinitesimal changes in x and t that it presupposes is certainly false. This is because claims about infinitesimal changes in x and t are fundamentally microscopic in nature, and we know that at the microscopic level, tiny changes in position and time look nothing like that implied by the above equation. Moreover, if we decide to squint, and ignore any implications the differential equation has in the microscopic realm, we completely lose our ability to use the differential equation to make *macroscopic* predictions – differential equations, after all, work by tracking the way in which infinitesimal changes 'add up' to produce non-infinitesimal changes.

One might try to avoid this inconvenience by not thinking of the derivative $\frac{dx}{dt}$ as a relation between infinitesimals, but rather as a relation between small changes in x and small changes in t – where these changes are small enough that they will be negligible on a macroscopic level, but large enough that we are safe enough in ignoring non-classical effects. Thus, $\frac{dx}{dt}(t_0)$ is replaced by $[x(t_0+\Delta t)-x(t_0)]/\Delta t$, and $\frac{d^2x}{dt^2}(t_0)$ by $[x(t_0+2\Delta t)-x(t_0)]/\Delta t^2$ in our differential equation, where Δt is small relative to the scale of the problem, but not so small that consideration of non-classical effects becomes necessary. But unfortunately, this move comes at a price – with this crass version of the differential equation, we will only be able to draw the crassest of conclusions. Our predictions will be less accurate at best, and at worst totally untrustworthy. (For instance, the kind of chaotic effects that were so interesting to us in the first place will be badly distorted, if not completely washed away.) We must therefore view our differential equation as a relation between infinitesimal changes in x and infinitesimal changes in t. But this involves attributing a type of microscopic detail to the system which is (i) empirically unwarranted, and (ii) almost certainly false, given our knowledge of the way the microscopic realm works.

Another way of trying to avoid this problem might be to try and eliminate talk of infinitesimals altogether, and simply view the differential equation above as a statement about the relationship between the values of x and $\frac{d^2x}{dt^2}$ at a given time t, where $\frac{d^2x}{dt^2}$ is thought of as an *intrinsic* property of the object at time t. This is all fine and well, but it does nothing to avoid the main problem. In all but the simplest of cases (i.e., except for the case in which an analytic solution is available by inspection), to get a solution out of a differential equation we need to perform a process of integration, which must, one way or another, amount to adding an integrand over tinier and tinier intervals of t. Thus, we need to make claims about the way in which tiny changes in a variable like x are connected with tiny changes in a variable like t. The same problems arise.

Another possible counterargument might be to say that there is nothing necessarily *microscopic* about relating infinitesimal changes in x and t. Just as we can consider the time evolution of a quantum system over infinitesimal, finite, or infinite times intervals; so too can we consider the time evolution of a *classical* system over infinitesimal, finite, or infinite times. Why should the fact that we are taking about infinitesimal time intervals mean that we are somehow doing something inappropriate if we maintain our classical language? (Of course, if we want an *exact* description of what happens over an infinitesimal time interval, our differential equation will not do the job – but if we want an exact description of what happens over a *finite* time interval, our differential equation will not do the job either; so the connection between the failure of exactness and the microscopic realm is therefore a red herring.)

But I don't think this counterargument works. We may use the differential equation in question to make claims about how incredibly tiny changes in x and t are related – and indeed, we must do so, in order to eventually get macroscopic solutions. But such claims border on the meaningless, insofar as they apply the physics of macroscopic time (t) and length (x) scales to microscopic time (Δx) and length (Δx) scales. This is precisely the 'unphysical' sort of reasoning that the proposal at hand wants to criticize – the projection of a given way of talking about the world outside regions of empirical support.

So whichever way we go, we are faced with the fact that macroscopic equations (such as the differential equation we have been considering) presuppose a picture of the microscopic which is unwarranted, and therefore 'unphysical' by the standards of our proposal. Thus, if the proposal is right, most calculations in physics involve unphysical argumentation.

(2.) Commitment to a law in a given regime can often lead, by innocuous argumentation, to commitments in quite distinct regimes.

Consider, for instance, Newton's law of gravitation. Fix a body of mass M (the sun for instance), and let us consider the force it exerts on various masses at a (fixed) distance r from its center of mass. Consider two bodies of mass m_1 and m_2; assume that their attraction to the mass M is given by Newton's law of gravitation; i.e., $F_i = GMm_i/r^2$. Place these bodies next to one another to form a single body of mass $m_1 + m_2$ – the force on this composite body is just the sum of the individual forces on m_1 and m_2, so that $F_{tot} = GMm_1/r^2 + GMm_2/r^2 = GM(m_1+m_2)/r^2$. Thus, if masses m_1 and m_2 are attracted (at distance r) to M by the usual force law, then the same will apply to $m_1 + m_2$.

Assuming that $m_1 > m_2$, we can also argue that a mass $m_1 - m_2$ will be attracted to M by the usual force law. In particular, split a mass m_1 into two masses $m_1 - m_2$ and m_2. From the facts that m_1 and m_2 obey the force law $F = GMm/r^2$, and that this expression for F is linear in m, it follows that the same force law applies to a point mass of magnitude $m_1 - m_2$.

Assume that we have convinced ourselves that for a range of masses $m \in (m_1, m_2)$, a point mass of magnitude m at a distance r from the body M will be attracted to it via the force law just given. Assume we are also committed to Newton's three laws (and in particular, to the law of action/ reaction.) Let S be the set of m with this property (i.e., the set of m such that a point mass of magnitude m at a distance r from the body M will be attracted to it via the usual force law.) Then S contains every real in (m_1, m_2), is closed under addition, and it also closed under subtraction. From this it quickly follows that S contains all positive reals.

Thus, from the fact that a certain range of masses at a distance r from the sun are attracted to the sun via the usual law of gravitation, it follows that *all* masses at a distance r from the sun are attracted to the sun via the usual law of gravitation (assuming Newton's Laws). So commitment to the law of gravitation in a certain range of cases can lead, via fairly innocuous argumentation, to the same law obtaining in a much larger range of cases. This kind of behaviour will not be uncommon in areas of physics whose laws are linear (e.g: classical electrodynamics, quantum mechanics). But again, it is precisely this sort of reasoning that our proposal wants to declare 'unphysical'.

What lessons can we draw from these two cases? We would like to think that we can commit ourselves to the laws of physics in certain regimes, without committing ourselves to any particular picture in other regimes. But (perhaps surprisingly) this is generally not possible. Mathematical formulations of laws have a way of 'spilling over' into unintended regimes – even into regimes in which we may well *expect* the laws in question to fail. Moreover, in the first case, we saw that it was *necessary* to exploit the physics of 'other regimes' in order to obtain information about the regime of interest.

The proposal we have been considering is that when a law is used outside its intended region of applicability, the resulting argument is 'unphysical'. But this means that an abundance of arguments that we took to be harmless are all of a sudden clouded with suspicion. Whenever we use differential equations – be they in classical mechanics, electrodynamics, or even quantum mechanics – we make claims about microscopic scales that are almost certainly false. Likewise, many of our laws have mathematical consequences in regimes in which we would surely be uncomfortable claiming definite knowledge. So, by the standards of this proposal, the most innocent physics turns out to be full of 'unphysicalilty'. I think we must conclude that we have gone wrong somewhere. While I still think that the proposal we have been considering here has an element of truth to it, it obviously needs some revision. Let us turn to this task.

4.4 INFERENTIAL RESTRICTIVENESS AGAIN.

4.4.1 Reasoning with Differential Equations.

I would like to consider a different (but related) analysis of 'unphysicality' in section 5. In order to do this, it will help to think awhile about the phenomenon discussed in the previous subsection, re-introducing the concept of an inferential restriction for help.

I argued that when a physicist uses a differential equation to make a prediction, he will be committed to the final result, but not necessarily to all of the claims he might make along the way. In particular, he will not be committed to claims involving values of Δx and Δt at small time scales, in spite of the fact that he must build upon such claims in order to make his predictions. Such intermediate claims are to be thought of as a means to an end – rather than as 'independently true' claims which are to be built upon, the way in which mathematicians build on lemmas and earlier results.

I think that this is most easily understood using the concept of an inferential restriction, as introduced in previous chapters. When the physicist numerically solves a differential equation, he is not committed to the truth of intermediate steps involving claims about tiny time scales. There is an inferential restriction which prohibits him from drawing conclusions about the way in which the world works at such scales. Just as the physicist is not allowed to remove a delta function from an integral, or say that 'the wire is straight' outside the context of a particular calculation, so too is he prohibited from making claims about infinitesimal time scales outside the context of a certain type of argument – namely, an argument in which, with the help of a differential equation, claims about infinitesimal space and time scales are 'added up' to draw conclusions about what happens at larger space and time scales.

This explains why it does not matter that, although physicists are willing only to endorse laws in certain regimes, those laws often presuppose (or entail) facts about other regimes. Specifically, this does not matter because inferential restrictions prohibit the physicist from drawing conclusions about the forbidden regimes, even though the physicist may have to (knowingly or unknowingly) use facts about such regimes en route to claims about the regime of interest. Inferential restrictiveness thereby stops the physicist from making claims to which he is not entitled.

I would like to dwell on this phenomenon a little longer. But first, it is interesting to note that a very similar (although not identical) phenomenon was also considered some time ago by Bridgman. On p. 262 of [11], Bridgman argues:

> *To make the question specific, suppose that we have established by experiment that the equation $d^2s/dt^2 = g$ describes the motion of any freely falling body. We then deduce from this equation by the logical steps of mathematics that the distance fallen by any body starting from rest is connected with the time of fall by the equation $s = \frac{1}{2}gt^2$. Is it superfluous to verify this result by measurement? I think a common answer could be that verification is quite superfluous, on the ground that the integrated form of the equation is already implied (whatever 'implied' may mean) in the differential equation from which we stated. I think a better answer would be that the answer depends on the physical operations which we have chosen to employ for our verification, first of the differential equation, and then of the integrated form. If the operations for these two verifications are such that both verifications may be made from the same set of readings in a note book, then we may say that the verification of the integrated equation is superfluous except in so far as it guards against our own blunders. But if the set of readings to verify the differential equation is not the same as made to verify the integrated equation, then verification is not superfluous.*

Bridgman's suggestion here is that we *cannot* infer $s = \frac{1}{2}gt^2$ from $\frac{d^2s}{dt^2} = g$, unless the kind of empirical evidence we would give for $s = \frac{1}{2}gt^2$ and the kind of empirical evidence we would give for $\frac{d^2s}{dt^2} = g$ somehow co-incide. Assuming this co-incidence does not obtain, Bridgman thinks (to use my language) that some sort of inferential restriction prohibits the physicist from inferring $s = \frac{1}{2}gt^2$ from $\frac{d^2s}{dt^2} = g$, in spite of the fact that the physicist is presumably allowed to infer basic facts about the acceleration of an object at any given time from $\frac{d^2s}{dt^2} = g$.

Of course, Bridgman's example is different from the one I have been considering. I have been wondering what stops the physicist from inferring facts about microscopic time scales from a macroscopic differential equation like $\frac{d^2s}{dt^2} = g$. Bridgman's worry goes roughly the other way around – he doesn't want us to be able to automatically infer $s = \frac{1}{2}gt^2$ from $\frac{d^2s}{dt^2} = g$ (exactly why is not important). In both cases, it is the imposition of inferential restrictions that saves the day; specifically, inferential restrictions help us when we (i) need to make a detailed mathematical assertion in order to do some calculation, but (ii) we do not want to take seriously *all* the mathematical consequences of our assertion. Inferential restrictions are what allow us to use such assertions, without overcommitting ourselves.

The worry someone might have, however, is that there is a type of irrationality in all this – that one has *no right* to restrict inferences, and that when we endorse a law we *must* thereby endorse all its mathematical consequences. In the next subsection I would like to continue my fight against this sort of worry. In the process, we will learn something about the nature of discovery and experimentation in physics, as well as garnering more evidence for both the Principle of Inferential Restrictiveness and the Deduction Failure Hypothesis.

4.4.2 Discovery in Physics.

The fact that a physicist cannot necessarily take seriously all the mathematical consequences of an assertion that *is* taken seriously should not be unfamiliar, given the content of the previous chapters – specifically, given the discussion of classical electrodynamics in Chapter 2, and the toy models (three light systems) discussed in both Chapters 2 and 3. As discussed there, with the help of the machinery of poset homomorphisms one can justify the use of claims (such as idealizations and mathematically incoherent concepts) not all of whose mathematical consequences can be trusted. Thus, the answer to the worry given at the end of the previous subsection is really already contained in what I have said in previous chapters. Nevertheless, it will be helpful to re-introduce some of these points in a slightly different way, in order to target the issues peculiar to the current chapter.

Let us consider a different example. Again, the example will involve a physicist making a claim far more specific than anything he would ordinarily commit to, in order to derive some macroscopic conclusion.

Assume that a 1 kg particle in one dimension is pushed around by some time dependent force. Imagine that a physicist wants to know the magnitude of this force, and the way in which it varies over time. He measures the position of the particle at various times; let us say that he concludes that $x(t) = t^2$ m. He then uses the classical equation $F = ma$ to conclude that $F(t) = 2$ N.

Now, if the physicist is asked 'at $t = 1$ s, is $x(t)$ *exactly* 1 m?', he will surely answer in the negative. Let us assume he admits a possible error of magnitude 2δ in his position measurements, so that although his answer to the previous question is negative, he is quite happy to affirm that, at any time t, $t^2 - \delta \leq x(t) \leq t^2 + \delta$.

The problem is that from $t^2 - \delta \leq x(t) \leq t^2 + \delta$, the physicist can infer *nothing* about the value of the force at any specific time. This is because for any real number r, there exists a function $f_r(t)$ such that:

1. For all t, $t^2 - \delta \leq f_r(t) \leq t^2 + \delta$, and

2. $f_r''(1) = r$.

In particular, if $r \leq 2$, let

$$f_r(t) = t^2 + \delta \ \mathrm{Exp}[-\frac{(2-r)}{2\delta}(t-1)^2],$$

and if $r \geq 2$, let

$$f_r(t) = t^2 - \delta \ \mathrm{Exp}[-\frac{(r-2)}{2\delta}(t-1)^2].$$

Our physicist finds himself in the following predicament: in order to be able to infer *anything* about the magnitude of the force, he must make a claim about $x(t)$ whose accuracy far exceeds anything to which he is entitled. Resigned to making such a claim about $x(t)$, and having made it, how will he treat it? He will trust *some* consequences of his claim that $x(t) = t^2$ m, but not others. This suggests that, associated with the claim $x(t) = t^2$ m, there will be a set of inferential restrictions that determines which conclusions may be drawn from $x(t) = t^2$ m, and which conclusions may not be drawn. This should not be surprising, given the discussion of the previous subsection.

Let us think about this situation a bit more. Imagine that our physicist makes the following claim: 'I have verified (or conjectured) that $x(t) = t^2$ m, and thus I conclude that the force on the particle is 2 N.' I take it that this is a sensible claim – indeed, it *must* be a sensible claim, because if it is not, then there is nothing sensible at all that the physicist can say in the given situation. Assuming then that it *is* a sensible thing for the physicist to say, let us think about what it could mean.

When the physicist says 'I have verified (or conjectured) that $x(t) = t^2$ m, and thus I conclude that $F = 2$ N', what he means is the following: he has performed some measurements, and on the basis of those measurements, he conjectures that in certain argumentative contexts, the assumption '$x(t) = t^2$ m' will lead to correct results. Moreover, he conjectures that the argumentative context in which one uses the identity $F(t) = mx''(t)$ to find the applied force is one of the permitted uses of the premise $x(t) = t^2$ m. Thus he conjectures that the conclusion $F(t) = 2$ N can be trusted.

We can break this process down a bit. The physicist makes a few position measurements, and all of those measurements are found to be consistent with $x(t) = t^2$ m, within a narrow margin of error. The physicist then conjectures that at *all* times, the position of the particle is given approximately by $x(t) = t^2$ m. This is not the conjecture I wish to dwell on, however. I simply mention it to distinguish it from the conjecture that I will really be interested in.

In making the conjecture that the position of the particle is approximately $x(t) = t^2$ m, our physicist is not done, because this conjecture is too weak for him to work with (for the reason just discussed). The stronger conjecture that $x(t)$ is exactly t^2 m would be too strong, however. So what the physicist does is makes an *additional* conjecture. He conjectures that he is in the sort of situation in which the claim '$x(t)$ is exactly t^2 m' can be used to draw *certain types* of conclusions – although not necessarily others – about the particle. In this case, the physicist conjectures that can infer that the applied force is approximately $F = mx''(t)$, where $x(t) = t^2$.

Of course, this is the sort of conjecture that can turn out to be wrong. We can imagine a world (or even an area of actual physics) in which positions oscillate so quickly and wildly that such arguments cannot be trusted. However, we are imagining (and *must* imagine) that the physicist is dealing with the sort of situation in which this typically does not happen; i.e., we imagine the physicist is in the sort of situation in which one *can* calculate approximate forces by twice differentiating the position function $x(t)$ obtained by smoothly interpolating a few data points. I am not suggesting here that, in virtue of being in such a fortunate situation, the physicist has simply gotten lucky – I am imagining that the physicist is acutely aware of the problems of differentiating an interpolated function, but recognizes *on the basis of experience* that the sort of situation he is in is likely to be one in which such differentiations can be trusted. Again – this is just a conjecture, and he could be wrong. But without such conjectures, our physicist would be lost. The important point then is that there is an additional conjecture here that (i) is necessary, and (ii) is not so much a conjecture that some 'fact' such as '$x(t)$ is exactly t^2 m' is true (because it is not), but is rather the conjecture that given the sort of situation he is in, certain arguments from the otherwise overly strong assumption that $x(t)$ is exactly t^2 m can be trusted.

The important point here is that although the physicist recognizes that the assumption $x(t)$ is exactly t^2 m works well in a variety of inferential contexts, the physicist will be *unwilling* to perform an induction and conclude that *all* arguments that use the assumption that $x(t)$ is exactly t^2 m can be trusted. For instance, from $x(t) = t^2$ one can infer that 'at time $t = 1$ s, the particle has traveled an integral number of meters.' Such a conclusion would be blatantly unwarranted.

It is important to distinguish the complicated set of conjectures that are and are not made when the physicist 'verifies' that '$x(t) = t^2$'. First, there is the conjecture involved when the physicist performs an interpolation on a finite set of data points, to arrive at '$x(t) = t^2$'. Tied up in this is the conjecture that '$x(t) = t^2$' works well in a variety of inferential contexts. But the conjecture that '$x(t) = t^2$' works well in *all* inferential contexts is *not* made – to do so would be to do bad physics.

Much the same can be said of the sort of situation we considered in Section 3. Imagine that theoretical considerations lead a physicist to postulate that a 1 kg particle in one dimension is being pushed about by some time dependent force $F(t) = \lambda \sin(\omega t)$; that is to say, that $x''(t) = \lambda\sin(\omega t)$. What can a physicist *mean* when he says that he has verified that $x''(t) = \lambda\sin(\omega t)$? I have argued that a differential equation of the form $x''(t) = \lambda\sin(\omega t)$ is ultimately a statement about the way infinitesimal changes in x and t are related, and it is difficult to imagine that the physicist has *any* empirical evidence that tells him *anything* about what happens on such time scales. What the physicist must mean when he says that he has verified that $x''(t) = \lambda\sin(\omega t)$, I claim, is that a reasonable range of mathematical arguments that use $x''(t) = \lambda\sin(\omega t)$ as a premise and draw macroscopic conclusions have been shown to be trustworthy.

But does this mean that the physicist then considers himself justified in performing an induction whereby he concludes that *all* mathematical arguments that use $x''(t) = \lambda\sin(\omega t)$ as a hypothesis have conclusions that can be trusted? Again, the answer is absolutely not – the physicist will want to remain completely agnostic about what *really* happens at time scales well beneath his empirical purview. The error of the inferential permissivist is to assume that such an induction is performed, but it seems clear to me that such an induction would have the physicist making claims to which he is utterly unentitled. The physicist has verified that $x''(t) = \lambda\sin(\omega t)$ works well as a hypothesis in *certain types* of arguments; but to say much more would be to do bad physics.

Justificatory questions concerning why *any* arguments from premises such as $x(t) = t^2$ m or $x''(t) = \lambda\sin(\omega t)$ can be trusted, given that not *all* such arguments can be trusted, can be addressed using the machinery of poset homomorphisms as in previous chapters. I shall quickly go through the construction for the later case.

Let L be any (fixed) set of laws describing a single particle system, which on macroscopic time scales yields $x''(t) \approx \lambda\sin(\omega t)$. What happens on other time scales is unimportant – all that matters is that the particle behave in the prescribed way macroscopically. Let T_1 be the set of claims that could be made of the one particle system that are consistent with L. For $p, q \in T_1$, say that $p \leq_1 q$ iff q may be deduced from p and L. As before, for $p, q \in T_1$, define $p =_1 q$ iff $p \leq_1 q$ and $q \leq_1 p$, and for any $p \in T_1$, let $[p]_1$ be the set of $q \in T_1$ such that $p =_1 q$. For $p, q \in T_1$, define $[p]_1 \leq_{[1]} [q]_1$ iff $p \leq_1 q$. Then $\leq_{[1]}$ defines a partial order on the set of equivalence classes $[p]_1$ with $p \in T_1$; call the resulting partially ordered set P_1. As before, P_1 encodes the implications that hold between assertions about the world from

God's perspective, i.e., in virtue of the physical laws L being as they are.

Let T_2 be the set of *approximate* claims that can be made about the way in which the position of the particle changes over *macroscopic* time scales, and which can be derived from the exact equation $x''(t) = \lambda \sin(\omega t)$. For $p, q \in T_2$, define $p \leq_2 q$ iff q is mathematically deducible from p and $x''(t) = \lambda \sin(\omega t)$. As before, define $p =_2 q$ iff $p \leq_2 q$ and $q \leq_2 p$. Also define equivalence classes $[p]_2$ for $p \in T_2$ and an induced partial ordering $\leq_{[2]}$ on these equivalence classes just as before, and call the resulting partially ordered set P_2. The partially ordered set P_2 encodes the mathematical implications that hold between *macroscopic* assertions about the particle, in virtue of the fact that $x''(t) = \lambda \sin(\omega t)$.

We then have the usual result:

Poset Homomorphism Result: $\forall p, q \in (T_1 \cap T_2)[([p]_2 \leq_{[2]} [q]_2) \rightarrow ([p]_1 \leq_{[1]} [q]_1)]$.

This result says that the hypothesis $x''(t) = \lambda \sin(\omega t)$ is correct, provided that it is only used to draw macroscopic, approximate conclusions about the way in which the position of the particle changes over time. As suggested in previous chapters, this poset homomorphism result should be viewed as a 'fact about the world' which justifies the inferentially restrictive use of $x''(t) = \lambda \sin(\omega t)$, while remaining agnostic about whether an inferentially permissive use of $x''(t) = \lambda \sin(\omega t)$ would be justified.

Let me summarize the main point of this section. Often, when a physicist says 'I have verified (or I conjecture) that ϕ', what is meant is that he has verified (or he conjectures) that there is a wide variety of mathematical arguments that use ϕ as a premise, and whose conclusions may be trusted. The physicist is, however, generally *unwilling* to perform an induction and argue that ϕ can be trusted in *all* inferential contexts. In many cases, such an induction would be obviously unwarranted (and in some cases, obviously false). In spite of the physicist's unwillingness to perform this induction, however, his use of ϕ in a limited set of inferential contexts can still be justified, in virtue of the relevant poset homomorphism.

Note that in virtue of the physicist's unwillingness to perform the induction just described, and conclude that ϕ can be trusted in *all* inferential contexts, we have more evidence for the Principle of Inferential Restrictiveness. Furthermore, in virtue of the fact that such an induction would require him to make claims outside the body of knowledge in question (as in the case of differential equations), we have more evidence for the Deduction Failure Principle.

4.5 UNPHYSICALITY.

In this section, I would like to build on the material of the previous section in order to give an analysis of 'unphysicality'.

When a physicist asserts that $x(t) = t^2$, or that $d^2s/dt^2 = g$, or when he asserts that Hamilton's equations are true or that Schrödinger's equation is true, his assertions are only to be used in a limited set of inferential contexts. Such assertions carry inferential restrictions with them. To violate such inferential restrictions is (at best) to make a claim without sufficient warrant, or even (at worst) to traffic in falsehood or meaninglessness. Which sorts of inferential contexts are permissible or not is something that is determined by the collective experience of the physics community. So much as in Chapter 2, the relevant inferential restrictions are determined by a process of conjecture and experience.

Of course, I did not say that when a physicist makes an assertion, the assertion is only to be trusted in a given regime, because as I have discussed, we often have to exploit what a law might say about regimes beyond our experience in order to draw conclusions about regimes within our experience. So there will be no blanket inferential restriction that prohibits *all* theorizing about distant regimes, even though there will surely be inferential restrictions that prohibit *some* such theorizing.

Given this, I propose the following analysis of 'unphysicality' – an argument is 'unphysical' if (i) it invokes some set S of laws of nature or descriptions of nature (e.g., it invokes Hamilton's equations, or the claim that a particle is an oscillator driven by an external periodic force $F(t) = \lambda \sin(\omega t)$), and (ii) it violates one of the inferential restrictions associated with some element of S.

So for instance, consider the case of the equation

$$\frac{d^2x}{dt^2} + \lambda(\frac{dx}{dt} - 1) = \Phi(x)P_T(t)$$

and its associated strange attractor. Some physicists might argue that there is an inferential restriction that prohibits us from beginning with the macroscopic differential equation just given, and concluding that the set of trajectories through phase space of the *physical* system has such and such a bizarre topological property, insofar as these later sorts of claims are fundamentally microscopic in nature. If this is right, the explanations of macroscopic properties of the system in terms of bizarre topological properties associated with a strange attractor would be regarded as unphysical. But note that counterarguments are possible here – in particular, Smith argues in [89] that we need not characterize strange attractors in terms of their bizarre topological properties; instead, we should characterize them as invariant sets under given transformations of phase space (cuttings, foldings, shearings, and so on). To finish the job, Smith would then have to argue that inferring macroscopic conclusions from the properties of the invariant set of some transformation of phase space associated with a classical dynamical system does not violate any reasonable inferential restriction.

I do not wish to adjudicate this debate – all I wish to do is note that it very much boils down to a question of (i) what our inferential restrictions are, and (ii) whether a given explanation can be paraphrased in such a way that these inferential restrictions are not violated.

Much the same goes for the three body problem. One might argue that the fancy solutions to the three body problem discussed earlier critically rely on the law of gravitation being microscopically valid, and that such orbits therefore violate the inferential restrictions that are naturally associated with the equation

$$F = G\frac{m_1 m_2}{r^2}.$$

One counterargument would be that circular, elliptical, and parabolic orbits are equally guilty of this. The question then boils down to exactly which sorts of microscopic uses of the inverse square are allowed, and which should be prohibited. If the sorts of experiences which permit physicists to use the microscopic inverse square law when discussing elliptical orbits (e.g., no rapid oscillatory changes in direction, etc.) seem to hold equally well of the strange orbits, then their physicality must be granted – if not, then their physicality could be plausibly denied. So the question of whether the strange orbits in question are physical seems to boil down to the same issues as before: (i) what our inferential restrictions are (with respect to the inverse square law) and (ii) whether such inferential restrictions are violated in postulating strange orbits.

One virtue of my proposal is that claims of 'unphysicality' will be highly fallible. In particular, as our experience expands, the set of inferential restrictions associated with any law will change. We will realize (as a product of theory and experiment) that some laws work well in new circumstances, but perhaps not so well as we thought in other areas. As our set of inferential restrictions changes, so too will our judgments of physicality. So, for instance, if we relax inferential restrictions, we may have to rescind earlier judgments of unphysicality. This seems to be what has happened with the acceptance of black holes by the physics community. On the other hand, if we enforce inferential restrictions, previously accepted arguments may be subsequently judged to be 'unphysical'. So for instance, invoking classical electrodynamics to explain the coherence of matter is now regarded as unphysical, because of the inferential restrictions that we now recognize must be placed on classical electrodynamics, as discussed in Chapter 2.

I think that inferential restrictiveness provides us with the right language in which to frame these debates. One might think that talk of energy, space and time regimes is sufficient, but it is not, as we quickly get mired down in the sorts of problems discussed in earlier sections.

I also think that the phenomena we have been examining provide us with further evidence for the Principle of Inferential Restrictiveness and the Deduction Failure Thesis. In particular, the fact that physicists recognize that some arguments might be unphysical – i.e., might violate some inferential restriction associated with some law or description of nature – is supporting evidence for the Principle of Inferential Restrictiveness, according to which physicists do not feel bound to accept the mathematical consequences of assertions that they accept. Likewise, the fact that some arguments *really are* unphysical (e.g., an argument in which facts about empirically inaccessible time scales are inferred from a differential equation) demonstrates the Deduction Failure Principle, according to which bodies of knowledge in physics are generally not closed under mathematical consequence.

4.6 CONCLUSIONS.

4.6.1 Summary.

I have spent the last three chapters looking closely at the methods of physics. In particular, I have considered three phenomena:

(1.) Physicists are often forced to *idealize* when they describe nature. In doing so, they produce descriptions of nature that they generally *know* to be false. In spite of this, physicists have an uncanny ability to reason with such falsifications in such a way that their conclusions turn out to be true.

(2.) Physicists are quite happy to reason with non-rigorous concepts, and do not feel that the use of such concepts compromises their calculations, so long as they obey certain rules associated with such concepts.

(3.) Physicists will sometimes dismiss – and be *right* in dismissing – mathematical consequences of assertions that they take seriously, arguing that the relevant arguments are in some ways *unphysical*.

These three phenomena provide us with ample evidence for the Principle of Inferential Restrictiveness, according to which physicists are perfectly comfortable using mathematics inferentially restrictively. That is, they provide evidence for the fact that physicists do not feel bound to accept the all the mathematical consequences of assertions that they accept.

In addition to the Principle of Inferential Restrictiveness, we have the Deduction Failure Thesis, according to which the *theories* that physicists produce are not necessarily closed under arbitrary mathematical deductions. In contrast with the Principle of Inferential Restrictiveness, this principle is not about what physicists feel comfortable doing, or how they like to use mathematics – this thesis is about the theories that they produce. We have seen in each of the three chapters that physicists produce bodies of knowledge whose mathematical consequences they do not take seriously. In some cases, not only *can* they not, but they *must* not be taken seriously, at the risk of producing physically false or even meaningless claims. This provides us with evidence for the Deduction Failure Principle.

So the big picture, roughly speaking is this – physicists do not feel committed simpliciter to the mathematical consequences of theories to which they are in turn committed. Because of this, they produce theories which are not closed under mathematical deduction, and whose deductive closure in fact often contains physically false sentences.

I have also argued that physicists are *justified* in being selective about which mathematical rules they use. This justification can be spelled out in term of poset homomorphisms. Roughly speaking, in virtue of facts about the world (which are learnt by a process of experience and conjecture), certain inferential structures embed into the inferential structure generated by the true laws of nature. This justifies the use of *certain* mathematical inferences, without justifying the use of *arbitrary* mathematical inferences – unless the poset homomorphism can be extended to a larger homomorphism, which in general it cannot.

In light of all this, I think that behind the sometimes schizophrenic mathematical methods of physics, one finds a coherent and defensible philosophical view of the role that mathematics plays within physical theories. Such methods are not just the symptom of laziness, confusion, or a bad philosophy of mathematics. That is not to say, of course, that *all* the unconventional mathematical methods of physics can be justified – just the ones I have considered. Other uses of mathematics may require a different justification, if they are justifiable at all.

Even though I have been selective in which uses of mathematics I have examined, I think that the uses I have focused on are representative of some of the more puzzling ways in which mathematics is used in science. In explaining some of the methods behind the madness of the physicist, we can therefore begin to grasp ways of using mathematics that will be invisible to the philosopher who focuses only on pure mathematics, or to the philosopher who mistakenly tries to force physics into the costume of pure mathematics.

4.6.2 Grandiose Conclusions: Physics versus Metaphysics.

I would like to close with a few more abstract, grandiose conclusions. (I use the word 'grandiose' in the most derogatory sense possible).

According to a long tradition that goes back at least to Aristotle (and probably earlier), metaphysics is the study of the fundamental truths regarding objects (beings) – or, as some have put it, the study of the 'true'. The principles of metaphysics are generally taken to have the property that *any* consequence (logical, or mathematical) of such a principle is also true.

So 'metaphysics' gives us one model for thinking about bodies of knowledge – i.e, as bodies of fundamental truths that may be used inferentially permissively to derive other, less fundamental truths.

For a rival epistemology, consider Cartwright. According to Cartwright, the laws of physics are true only in models, which are (for Cartwright) fictional scenarios. Thus, Cartwright thinks that the laws of physics do *not* describe reality – they describe a fictional reality which, although similar to ours, is nevertheless distinct from ours.

Between these view lies the epistemology that I would like to maintain. The laws of physics as we know them are not valid taken inferentially permissively. But nor are they fiction. The world is built in such a way that from certain bodies of knowledge, certain types of valid mathematical deductions take us from true claims about reality to other true claims about reality. But not all such otherwise valid mathematical deductions do this. This is a fact about the world that physicists must struggle to get their hands around – to learn which deductions are good in which contexts, and which are not. The important point is that what I have just described concerning the efficacy of mathematical arguments is a fact about the *world*, and not just a fact about some work of creative fiction. So: physics is neither metaphysics nor fiction, but lies between the two.

On p. 357 of [83], Schwartz complains:

> *The mathematician turns the scientist's theoretical assumptions; i.e., convenient points of analytical emphasis, into axioms, then takes these axioms literally. This brings with it the danger that he may also persuade the scientist to take these axioms literally. The question, central to the scientific investigation but intensely disturbing in the mathematical context – what happens to all this if the axioms are relaxed? – is thereby put into shadow.*

To put my words in Schwartz's mouth, Schwartz is accusing a certain type of mathematician of confusing physics with metaphysics. At the other extreme, social constructivists and others often confuse science with fable, myth, patriarchal tools of power, and so on. Both viewpoints are quite erroneous, and entirely avoidable.

5.0 PROLOGUE TO PART 2

5.1 INTRODUCTION

Thus far, I have argued that the sorts of theoretical strategies that physicists often use tend to lead them to bodies of knowledge that are not closed under arbitrary mathematical inferences. This, of course, is just to repeat the combination of the Deduction Failure Thesis and the Principle of Inferential Restrictiveness.

One might wonder, however, how much of this is due to the fact that we have not uncovered the final laws of physics, and that we do not generally have exact microscopic descriptions of systems that we seek to understand. Perhaps if we knew the exact laws of physics and the exact descriptions of physical system of interest, we would no longer be confronted by the sorts of problems identified in Chapter 1, where I argued that reasoning with idealizations tends to require inferential restrictions. Likewise, one might imagine that the final laws of physics are statable in a vocabulary that does not rely on incoherent mathematical concepts, thereby avoiding the problems identified in Chapter 2. Finally, one might imagine that presented with the final laws of physics, and an exact description of any system of interest, we would never really have to worry about taking any mathematical argument 'too seriously'. In such a situation, we could do physics simply by blindly proving theorems, as if we were pure mathematicians. The sorts of problems identified in Chapter 3 – whose solutions required the adoption of inferential restrictions – would therefore not haunt us.

Thus, it might seem that the sorts of problems that have been identified in Part 1 are, at best, problems that arise from our own ignorance. There is a point of view – the God's eye point of view – from which these problems do not arise. Amongst all the other records God keeps, He has exact mathematical representations of all physical systems, and knowledge of the exact equations that govern the evolution of those systems in time. With this knowledge, the problems of Part 1 can be avoided.

However, this point of view presupposes that there *is* an ultimate mathematical description of reality (initial conditions plus laws) that could, in principle, be used to derive more complicated facts about reality. In part 1, I called this the Representation Thesis:

The Representation Thesis: The state of any physical system at any time, as well as all the fundamental laws of nature, must be able to be captured perfectly by some set of mathematical assertions or equations.

The Representation Thesis seems very natural. Moreover, it seems to accord with the methods of physics – when studying a physical system, the first thing a physicist will often do is write down some mathematical representation of the system. Is there any reason to think that this will be any more difficult for God than it is for us? If not, the Representation Thesis will seem compelling.

But I think that the argument sketch just given for the Representation Thesis is highly superficial. When we write down a mathematical representation of a physical system, along with its laws, we idealize and simplify in many ways. The fact that we have become good at writing down imperfect mathematical representations of physical systems and their laws does not entail that we could, if we were as smart as God, also give *perfect* mathematical representations. Nor, as far as I can tell, does our success at writing down better and better mathematical representations of a system and its laws count as good *inductive* grounds for thinking that it is in principle possible to give a perfect mathematical representation. The possibility of a sequence of events does not necessarily give us grounds for believing in the possibility of their 'limit' – for instance, the possibility of a recursive enumeration of all true sentences of arithmetic with less than n symbols, for any n, does not give us any sort of grounds for believing in the possibility of a recursive enumeration of *all* true sentences of arithmetic. Yet much of the attraction of the Representation Thesis comes, I think, from precisely this sort of fallacious move. (We will look at this and other motivations for the Representation Thesis in subsequent chapters.)

In fact, I would like to argue in Part 2 that the Representation Thesis is *false*. Specifically, I will argue for the following:

The Representation Failure Thesis: We cannot take for granted that the world is representable in mathematical terms – that is to say, we cannot take for granted that the state of the world, together with all its fundamental laws, can be captured perfectly by some set of mathematical assertions or equations. In fact, it is more likely that the world is *not* representable in mathematical terms (in this sense) than that it is.

Note that the Representation Failure Thesis states more than just the denial of the Representation Thesis. Not only does the Representation Failure Thesis state that it is *possible* for the world to fail to be mathematically representable – it also states that it is *more likely* that this is so than not.

I shall present two arguments for the Representation Failure Thesis – the first in Chapter 6, and the second in Chapters 7–8. I shall outline the rough strategy of each of these arguments now.

5.2 OUTLINE OF PART 2

5.2.1 Chapter 6: Is Mathematics the Language of Nature?

One reason for believing in the Representation Thesis is that it seems natural to think that the world is isomorphic to some sort of mathematical structure. Physics, then, is just the theory of this structure. But must this picture be taken for granted?

I would like to argue that the answer here is No. Specifically, I shall argue that it is possible for the world to fail to be isomorphic to any sort of mathematical structure. This argument will use the set-theoretic technique of forcing, together with the assumptions (i) that every logically consistent state of affairs is possible (including those described by sentences that are not first-order), and (ii) that the referent of the term $\mathbb{R}$ is the same in all possible worlds. With all this we will be able to explicitly describe a possible world that fails to be isomorphic to any mathematical structure.

In fact, I shall go on to argue that *most* possible worlds (in a sense to be defined) fail to be isomorphic to any mathematical structure. From this, the Representation Failure Thesis

Follows.

So, in fact, the mathematical representability of the whole world is a very special state of affairs. Of course, this does not contradict the fact that we may often be able to use partial or approximate mathematical representations to our profit, even when the world as a whole fails to be mathematically representable. But we *cannot* blindly assume that, from God's point of view, some set of mathematical assertions perfectly captures reality. The use of mathematics in most cases is therefore restricted only to approximate, or partial theories.

Faced with the fact that physicists tend to produce theories that violate the Deduction Thesis, we *cannot* simply blame our ignorance – that is, we cannot simply blame the fact that we do not see through God's eyes. Rather, I would like to suggest that the fact that physicists tend to produce theories that violate the Deduction Thesis should be viewed a symptom of a deeper incongruity between mathematics and physics. The fact that most possible worlds fail to be isomorphic to any mathematical structure is equally a symptom of such an incongruity.

5.2.2 Chapters 7–8: Is Mathematics Unreasonably Effective in Physics?

One might wonder what implications the previous chapter, as well as Part 1, have on the doctrine originally presented by Wigner in [101], according to which there is something 'unreasonable' about the effectiveness of mathematics in the sciences (and especially in physics.) It will be worth discussing the matter in some detail. I will divide my discussion into two chapters, Chapters 7 and 8. The issues discussed in Chapter 7 do not revolve around the Representation Failure Thesis, and thus are separated from the issues discussed in Chapter 8, which do revolve around the Representation Failure Thesis. (In this sense, Chapter 7 can be regarded as a digression from the main line of argument of Part 2 – even though the discussion of Chapter 7 will help to provide a context for the arguments of Chapter 8.)

A great many things can be meant by the claim that mathematics is 'unreasonably effective' in physics. Chapter 7 will revolve around a discussion of two particular ways in which this claim might be fleshed out. First, one might consider there to be something surprising about the fact that mathematicians have an uncanny knack for discovering and studying structures that only *subsequently* turn out to be useful in physics. Second, one might think that there is something peculiar about the fact that the mathematicians' *aesthetic* sense – which would appear to have very little to do with the physical world – turns out to be an extraordinary guide in sniffing out empirically successful physical theories. In Chapter 7 I shall examine both of these claims closely, and shall argue that in neither sense do we have any sort of reason to be surprised at the efficacy of mathematics in physics.

In Chapter 8, I would like to continue the examination of possible senses in which mathematics might be said to be 'unreasonably effective' in physics. There, however, I would like to argue that there *is* a sense in which the efficacy of mathematics in physics could turn out to be surprising. Specifically, I would like to argue that *if* it turned out that one could describe any actual system (both initial conditions and laws) in purely mathematical terms, then this *would* turn out to be genuinely surprising. Perhaps this point can be made just by focusing on the material of Chapter 6 – we have no good reason to expect that the world is isomorphic to a mathematical structure, so if it turns out that it is, we have a genuine cause for surprise, in much the same way that we have a cause for surprise given any low probability event. However, in Chapter 8, I would like to add two important features to this argument, to help strengthen it.

First (and most importantly for the discussion of the Representation Failure Thesis) I want to argue that even when the world *is* isomorphic to a mathematical structure, it does not follow that there must be some set of equations or mathematical assertions that captures the laws true of the world. It is possible for the world to be isomorphic to a mathematical structure, and yet for mathematics to fail to be the 'language of nature' (in a sense to be precisely defined, but different from that presented in Chapter 6.) In fact, the possible worlds which are isomorphic to a mathematical structure, and in which the language of nature *is* mathematics (in the sense defined in Chapter 8) are vastly outnumbered by the set of possible worlds isomorphic to a mathematical structure, in which the language of nature is *not* mathematics in this sense. Thus, I shall argue that the set of possible worlds in which mathematics is the language of nature in this richer sense of the word turns out to be a minority within a minority. That the actual world should turn out to be such a world would surely be a real surprise.

Second, I wish to look at the notion of a 'surprise' a little more closely. It is not immediately clear that all low probability events are surprising. I shall examine what it means to say that an event is surprising, and argue that the actual world turning out to be a world in which the language of nature is mathematics (in the richer sense of Chapter 8) is indeed a bona fide surprise. This helps to nail home the depth of the 'incongruity' of mathematics and physics mentioned in Chapter 1. Not only *cannot* one assume that the expressive capacities of mathematics 'match up' with physical reality – but moreover, it would be surprising if they *did* 'match up'. In this sense, the incongruity between mathematics and physics is strong indeed.

6.0 IS MATHEMATICS THE LANGUAGE OF NATURE?

6.1 INTRODUCTION.

In this section, I shall present my first argument for the Representation Failure Thesis. The strategy is as summarized in the previous chapter.

To begin, let us consider the thesis that 'the language of nature is mathematics'. This thesis is not only often cited by physicists, but also occupies a fundamental role in many philosophers' ways of thinking about the workings of nature. In fact, for most thinkers who have been raised with a modern scientific temperament, the thesis has become somewhat of a commonplace – stated so often, and by intellectuals of such stature, that one wonders how any sort of discourse in physics could be possible that did not presuppose it.

Most would say that the thesis goes back at least to Pythagoras. See, for instance, Aristotle's Metaphysics [2] xii. 6; 1080 b 16:

> *"The Pythagoreans say that there is but one number, the mathematical, but things of sense are not separated from this, for they are composed of it; indeed, they construct the whole heaven out of numbers . . . "*

Centuries later, we find Kepler saying (see [18]):

> *"The chief aim of all investigations of the external world should be to discover the rational order and harmony which has been imposed on it by God and which He revealed to us in the language of mathematics."*

Consider also the following quote from Galileo's *Assayer* (cited on p. 64 of [55]):

> *"Philosophy is written in this grand book of the universe, which stands continually open to our gaze. But the book cannot be understood unless one first learns to comprehend the language and to read the alphabet of which it is composed. It is written in the language of mathematics, and its characters are triangles, circles, and other geometric figures, without which it is humanly impossible to understand a single word of it; without these, one wanders in a dark labyrinth."*

From no less than Gauss we have (see [75]):

> *"God is a mathematician."*

And finally, to use an example from modern times, Einstein tells us (in his Herbert Spenser Lecture; see pp. 270-276 of [30]):

> *"Our experience hitherto justifies us in believing that nature is the realisation of the simplest conceivable mathematical ideas. I am convinced that we can discover by means of purely mathematical constructions the concepts and laws connecting them with each other, which furnish the key to the understanding of natural phenomena. ... In a certain sense, therefore, I hold it true that pure thought can grasp reality, as the ancients dreamed."*

No doubt, Pythagoras, Kepler, Galileo, Gauss and Einstein all had different things in mind when they praised mathematics in their various ways. At a very minimum, each seemed to think that all the processes of nature could be completely described in mathematical terms. But in addition to this, I think that each also entertained a slightly more ambitious view, according to which the natural world *itself* (and not just the processes that compose it) could be taken to be isomorphic to some giant mathematical structure.[1] The temporal evolution of this structure is governed, in turn, by precise mathematical equations.

There are a few different ways of making this more ambitious idea totally explicit. One could think of the world[2] *at any particular time* as given by some giant mathematical structure. One could then postulate that the various giant mathematical structures at different times can be connected together in some lawlike, mathematical way. This gives us a first way in which we can think of the natural world as isomorphic to some giant mathematical structure – specifically, a giant mathematical structure that changes over time.

Alternatively, one could just think of *trans-temporal* nature – i.e., the world as it exists *through* time, and not just the world at a *particular* time – as given by an even more enormous mathematical structure. Traditionally, this has involved thinking of the world as a giant, four-dimensional structure. This gives us a second way of thinking of the natural world as isomorphic to some mathematical structure.

For the sake of simplicity, I shall focus on this later formulation of the idea that I have

[1]In fact, for the Pythagoreans, the world is not merely *isomorphic* to such a mathematical structure – it is *identical* with such a structure.

[2]By the 'world' I here mean one particular world – generally the actual world – as opposed to the collection of all worlds, both possible and actual.

attributed to Pythagoras, Kepler, Galileo, Gauss and Einstein.[3] This thesis – *that trans-temporal reality is isomorphic to some enormous mathematical structure* – is what I shall mean by the claim that 'the language of nature is mathematics'.[4]

In spite of its widespread acceptance, there are a few reservations one might have about this thesis. First, one might worry about to what extent 'higher order' phenomena must necessarily be mathematical in nature. For instance – does it really follow from the fact that the language of micro-physics is mathematics, that human interactions or political parties can be described and understood in purely mathematical ways too? Anyone with anti-reductionist sympathies would surely be uncomfortable here.

One might also worry that there are facts about sense-experience that necessarily resist mathematical analysis. I know what the color red looks like, but is it really possible that all there is to know about redness can be captured in some mathematical structure? When one thinks of the world as given by some giant mathematical structure, it is surely conceded that many phenomenological properties of the world (for example, most secondary qualities) will have been abstracted away. But if so much is abstracted away, what right do we then have to say that the language of nature is mathematics?

For a further objection, recall that Henri Bergson argued in [10] that there were properties of objects that were in principle incapable of mathematical analysis (and in fact, incapable of any sort of *conceptual* analysis.) Bergson felt that the motion of a body was an example of this, insofar as any mathematical representation of the body that left us with nothing more than a union of time-slices necessarily abstracted away the 'mobility' of the body. Again, this might be taken to challenge the notion that an exhaustive discussion of nature can occur in purely mathematical terms.

Finally, one might worry that any mathematical representation of reality must make idealized assumptions that need not be true. Objects will be treated as point particles, or continua, or wavefunctions in mathematically convenient ways. Is it really obvious that these

[3]By focusing on this second alternative, I avoid issues about the nature of the lawlike connections that connect the giant time-slices of nature.

[4]My purpose here is not a detailed defense of the historical claim that the Pythagoreans, Galileo, Kepler, Gauss and Einstein held one particular version of this view, as opposed to another. I simply wish to suggest that the thesis in question has, in one form or another, exerted a great influence on Western science. It is the thesis itself, and not its pedigree or historical development, that will be the real object of attention here.

idealizations are harmless? If they are not harmless, or if we do not know whether they are harmless, our ability to confidently declare that the language of nature is mathematics might also be thought to be compromised.

Each of these objections deserves careful analysis. However, for the sake of the present chapter, I wish to put all of them to the side. Let us forget about 'higher-order' states of affairs, phenomenological properties and secondary qualities. Let us not worry about whether a mathematical representation can truly capture the 'mobility' or 'duration' of objections in motion. Furthermore, let us blindly assume that the world is made of nothing more than point particles floating around in space. Given all these concessions, I now ask: *must the language of nature be mathematics?* This question will occupy us for much of the chapter.

Spelt out in more detail, my question is as follows: let us assume that the world is nothing more than a collection of point particles moving around in space, and that the only things we care about are the spatio-temporal trajectories of these particles. Does it follow that the world must be isomorphic to some gigantic mathematical structure? I would like to argue that the answer is No.

In the next section, I will give my main argument for this claim. But first I would like to consider an argument that one can imagine a mathematician giving *against* my main thesis – that is, an argument the mathematician might give in *support* of the claim that the world must be isomorphic to some giant mathematical structure.

The argument goes like this: label all the points of space with the elements of some set S, and associate with each point particle P a function $f_P : \mathbb{R} \to P$ representing the trajectory of the particle through time in the obvious way.[5] Also, associate with the set S some other sort of mathematical structure that encodes the spatial relations that hold amongst the different elements of S (perhaps a metric, or a topology). For the sake of definiteness, let us assume that a metric ρ has been associated with S. We may then mathematically represent the world by the structure $<S, \rho, F>$, where $F = \{f_P : P \text{ is an existing point particle}\}$. The mathematical structure $<S, \rho, F>$ is then 'isomorphic' to the physical world, in the sense that every possible relation between particles and/or points of space may be captured by

[5]It is assumed here that time is structured like the reals $\mathbb{R}$. Of course, one can formulate a more abstract conception of the trajectory of a particle that does not depend on this assumption, but because nothing that follows is affected by this, I do not opt for such generality here.

some purely mathematical statement about the structure $<S, \rho, F>$. It follows that, in the sense I have described, mathematics may be taken to be the language of nature.

But let us consider this proposal more closely. The mathematician begins by asking us to 'label the points of space with the elements of some set S'. What does he mean by this? Presumably, he means that we should give a name to each point in space, and then throw all these names into a big set. But how do we know that this is possible? We can certainly imagine giving names to a finite handful of points in space. But what is supposed to happen after that? How do we know that all the points of space can be 'enumerated' in the way that the mathematician will require? And how do we know that we won't run out of names?

Certainly, one can invoke abstract principles such as Zorn's lemma to try and answer such questions. But these principles are purely mathematical principles, and the only way they can have any bearing on the question of whether or not one can associate a unique name with each point of space, is if one assumes from the beginning that space is some sort of mathematical structure. In this way, such fancy maneuvers simply beg the question.

Let me distinguish my difficulty with the mathematician's demand that we 'label the points of space with the elements of some set S' from a couple of difficulties with which it could be confused. My objection is *not* that, for all we know, the points of space may form a 'class', rather than a 'set'. Certainly, that would be one way in which the mathematician's demand could turn out to be unreasonable. But this objection presupposes that one *can* successfully label all the points of space with the elements of some sort of structure S – but that S happens to be a class, rather than a set. My point is that it is not obvious why *any* attempt to label the points of space with *any* sort of mathematical structure S should be successful.

Nor is my objection that there may not be any 'constructive' way of labeling the points of space with elements of some set S. My objection is not that the mathematician has given us no explicit algorithm for carrying out his demand, but rather that the mathematician has given us no reason for thinking that his demand could be carried out in *any* way, constructive or otherwise.

There is no reason to think that the mathematician's demand that we 'label the points of space with the elements of some set S' can be met, *without* presupposing that space *already* has some sort of mathematical structure. (I shall say more about this in section §4.) Certainly, the mathematician's assertion that 'one can label the points of space with the elements of some set S' is perfectly acceptable as an expression of faith in the mathematical structure of the universe. It cannot, however, count as a *justification* of such faith. In fact, it is precisely this sort of faith that I shall try to undermine in the next section.

6.2 THE MAIN ARGUMENT.

In this section, I shall argue that it is possible that the world not be isomorphic to any mathematical structure, and thus that the language of nature need not be mathematics (in the sense defined in the previous section). I shall first present a rough statement of the idea behind my argument, and shall then carefully go through the details.

Consider a single, solitary particle living in a 3 dimensional space. At time $t = 0$, imagine erecting a plane at $x = q$ (for some rational q), and asking whether the particle lies to left or right of the plane. (We assume at $t = 0$ that $x(t)$ is irrational, so it cannot lie 'on' the plane.) This experiment will yield a yes or no answer. We can furthermore imagine the result we would have obtained had we placed the plane at $x = q'$, for some rational $q \neq q'$. In fact, given an enumeration $q_1, q_2, q_3, ...$ of the rationals, we can imagine the sequence of yes/no responses that would come from performing the given experiment with the plane at $x = q_i$ and at $t = 0$. I shall argue that there are logically possible outcomes for this sequence of yes/no responses which are such that, if they occur, then the universe consisting of this solitary particle defies mathematical description. That is to say, there are logically possible outcomes to this countable sequence of counterfactual experiments which are such that, if they occur, then the language of nature is not mathematics.

Let us now go through the details. The argument will require two premises - one about the nature of possibility, and another about the nature of mathematics. The premises are as follows:

Modal Premise: Every logically consistent state of affairs is possible.

Mathematical Premise: The referent of the term $\mathbb{R}$ is the same in all possible worlds.

(The term $\mathbb{R}$ refers to the set of reals – i.e., the set of Dedekind cuts of the rational numbers.[6] Likewise, the terms $\mathbb{N}$ and $\mathbb{Q}$ will refer to the set of natural numbers and rational numbers.)

The second premise is a little less familiar than the first, and so I shall discuss it briefly before commencing my main argument. Let us imagine someone in a possible world who has constructed (or obtained by non-constructive means) some real number r. Imagine this person asserting that r is a real number. The content of the Mathematical Premise is that any other person in any other possible world will be able to recognize this claim as both meaningful and true. The meaningfulness and truth of claims of the form '$r \in \mathbb{R}$' are independent of any particular possible world. So, if a claim of the form '$r \in \mathbb{R}$' is meaningful and true in *some* possible world, then it is *necessarily* meaningful and true.

In addition, we assume that the set of natural numbers in the transworld referent of $\mathbb{R}$ are all standard – i.e., that the transworld referent of $\mathbb{R}$ contains no nonstandard natural numbers. (We take this to be a 'subpremise' of the Mathematical Premise, or an implicit Second Mathematical Premise.)

Of course, either of the two premises just listed can be denied. In a later section (specifically, §4), I will discuss the extent to which the conclusions of the present section can be avoided by denying either the Modal Premise or the Mathematical Premise. For now, however, let us press on, asumming both the Modal and Mathematical Premises to be correct.

My argument rests on a fact from set theory. Fix a model $V \models ZFC$. Let $\mathbb{R}^V, \mathbb{Q}^V$ and $\mathbb{N}^V$ denote the set of reals, rationals, and natural numbers contained in V. In addition, if T is a a set of sentences, define $Con(T)$ to be the proposition that says that T is consistent (in a sense to be clarified below). One may then prove the following theorem:

Main Theorem: $Con(ZFC) \rightarrow Con(ZFC+\text{there is a Dedekind cut of } \mathbb{Q}^V \text{ not contained in } \mathbb{R}^V)$

The sentence 'there is a Dedekind cut of $\mathbb{Q}^V$ not contained in $\mathbb{R}^V$' is to be thought of as

[6] By a 'Dedekind cut' of the set of rationals $\mathbb{Q}$, we mean a subset S of $\mathbb{Q}$ such that:

(i) $S \neq \emptyset$ and $S \neq \mathbb{Q}$,
(ii) $\forall p, q \in \mathbb{Q}[q \in S \text{ and } p < q \rightarrow p \in S]$, and
(iii) S contains no largest element.

For a discussion of Dedekind cuts, see [80].

an existential quantification over the conjunction of an infinite set of sentences, specifically:

$$\text{‘there is a Dedekind cut of } \mathbb{Q}^V \text{ not contained in } \mathbb{R}^V\text{’} \leftrightarrow (\exists x)\ [s(x) \wedge \bigwedge_{r \in \mathbb{R}^V} t_r(x)]$$

where $s(x)$ is the formula ‘x is a Dedekind cut of $\mathbb{Q}^V$’, and for each $r \in \mathbb{R}^V$, $t_r(x)$ is the formula ‘$x \neq r$’.

The sentence ‘x is a Dedekind cut of $\mathbb{Q}^V$’ should, in turn, be thought of as the infinitary sentence:

$$(\forall d)[(d \in x \rightarrow \bigvee_{q \in \mathbb{Q}^V} d = q) \wedge \Phi(x)],$$

where $\Phi(x)$ consists of the usual conditions for a Dedekind cut, given in the previous footnote.

Once we introduce infinite conjunctions and disjunctions, we need to re-specify what is meant by ‘consistency’. To say that a set of sentences is consistent is to say that no contradiction may be deduced from them. But what are the rules of inference relative to which we are to make this judgment?

To the usual logical axioms and rules of inference of the predicate calculus, we adjoin the logical axiom:

$$(\bigwedge_{\alpha<\beta} \Phi_\alpha) \rightarrow \Phi_\gamma \quad (\text{for each } \beta \text{ and } \gamma < \beta)$$

and the rule of inference:

$$\text{if } \Phi_\alpha \text{ for all } \alpha < \beta, \text{ infer } \bigwedge_{\alpha<\beta} \Phi_\alpha.$$

We must then allow for the possibility of proofs of arbitrary (infinite) length. For further discussions of proofs in infinitary logic, see, for example, Dickman [25], and §5 of Chapter III of Barwise [7].

To say that a set of sentences S involving infinitary connectives is consistent is then just to say that no sentence of the form $\Phi \wedge \neg\Phi$ can be derived from S, using the logical axioms and rules of inference just described.[7]

The Main Theorem tells us that if ZFC is consistent, then so is ZFC together with the assumption that there is a Dedekind cut of $\mathbb{Q}^V$ not in $\mathbb{R}^V$. I shall present a proof of this theorem in Appendix B. Let us press on, however, assuming this result to be true.

[7]The reader versed in such things will note that I have just described the infinitary logic $\mathcal{L}_{\infty,\omega}$.

Fix a model $V \models ZFC$ for which $\mathbb{R}^V$ is the unique trans-world referent of $\mathbb{R}$ referred to in the Mathematical Premise. Now consider the following proposition (call it X):

> X: 'Space-time has a 4-dimensional Euclidean geometry, and there exists a point particle P whose trajectory $\vec{f}(t)$ satisfies $f_x(0) = r$, where f_x is the x co-ordinate of $\vec{f}$, and r is a Dedekind cut of $\mathbb{Q}^V$ not contained in $\mathbb{R}^V$.'

One might worry that, at this stage, a claim like $f_x(0) = r$, as it appears in X, is meaningless. Specfically, one might worry that, even having said that space-time has a metric structure isomorphic to $\mathbb{R}^4$, one still has to specify an origin and a set of orthogonal vectors $\hat{x}, \hat{y}, \hat{z}$ with which to define a co-ordinate system, before a claim like $f_x(0) = r$ even makes sense. These details are easily filled in. Let us specify in X that space-time is to contain at least four uniquely identifiable and simultaneous events E, E_1, E_2 and E_3. Then one can specify that E is to represent the origin of our co-ordinate system, and use the vectors from E to E_1, E to E_2, and E to E_3 (assuming that they are not coplanar, and that no pair is collinear) to define an orthonormal co-ordinate system $\hat{x}, \hat{y}$ and $\hat{z}$. One can then specify that $f_x(0) = r$ relative to this co-ordinate system.

In addition, one might similarly wonder whether at this stage a claim like 'r is a Dedekind cut of $\mathbb{Q}^V$ not contained in $\mathbb{R}^V$' has any meaning, given that all we know about r is that it is a point in a 4-dimensional Euclidean space. Let us therefore explain what we take the meaning of 'r is a Dedekind cut of $\mathbb{Q}^V$ not contained in $\mathbb{R}^V$' to be. When we say that space-time has a 4-dimensional Euclidean geometry, we mean that space-time is given by a structure $\bar{\mathbb{R}}^4$, where $\bar{\mathbb{R}}$ is the set of Dedekind cuts, with the usual metric, of some ordered field $\mathbb{F}$. Now there exists a unique embedding $i : \mathbb{Q} \to \mathbb{F}$ of the (standard) rationals $\mathbb{Q} = \mathbb{Q}^V$ into $\mathbb{F}$. We then say that 'r is a Dedekind cut of $\mathbb{Q}^V$ not contained in $\mathbb{R}^V$' just in case $i^{-1}(r \cap i(\mathbb{Q}^V)) \notin \mathbb{R}^V$. Note that the structure $\bar{\mathbb{R}}$ is *rigid*, i.e., admits no automorphisms other than the identity, and so the truth value of 'r is a Dedekind cut of $\mathbb{Q}^V$ not contained in $\mathbb{R}^V$' is unambiguously determined by the criterion just outlined.

The Main Theorem tells us that the proposition X is logically consistent. Therefore, using the Modal Premise, X describes a possible state of affairs. Let w be a possible world in which X is true.

I claim that there there is no mathematical structure that can be used to represent the trajectory of P in w, without demanding that the reference of $\mathbb{R}$ in w be different from $\mathbb{R}^V$; i.e., without demanding that there is a Dedekind cut of $\mathbb{Q}$ not to be found in $\mathbb{R}^V$. In other words, the existence of a mathematical structure representing the possible world w runs afoul of the Mathematical Premise. Let us consider why.

Obviously, we cannot represent space-time with the mathematical structure $\mathbb{R}^4$, because then the trajectory of P would have to be represented by some $\vec{f} : \mathbb{R} \to \mathbb{R}^3$, where $\vec{f}(t)$ is the position of the particle at time t. This would force us to conclude that $f_x(0) \in \mathbb{R}$, which would be a contradiction. But it does not follow from this fact alone that the trajectory of P is mathematically ineffable. Perhaps we can use some unconventional mathematical trick to represent the trajectory of P. It does not follow from the fact that *one* way of representing P fails, that *all* ways must also fail. Perhaps, if we allow ourselves to exercise some mathematical imagination, what appears to be mathematically ineffable might not be mathematically ineffable after all.

I shall argue, however, that if the trajectory of the point particle P can be mathematically represented in *any* way – by *any* sort of mathematical structure – then one can construct, from such a structure, a Dedekind cut of $\mathbb{Q}^V$ not in $\mathbb{R}^V$. This will then contradict the Mathematical Premise, and we will be able to conclude that, in w, the language of nature is not mathematics. Let us see why.

A basic requirement of any mathematical representation of the trajectory of a particle is that, for any proposition p about the relative positions of objects, there should exist some purely mathematical claim M_p such that p holds if and only if M_p holds. As a simple example of this, for any $q \in \mathbb{Q}^V$, let the proposition p_q be as follows:

$$p_q : \quad \text{‘at time } t = 0\text{, the particle } P \text{ is to the left of the screen determined by } x = q.\text{’}$$

This is not a purely mathematical claim, as it involves the ordinary language expression ‘to the left of’. If one takes trajectories to be represented by functions $f : \mathbb{R}^V \to \mathbb{R}^V$, however, then one can construct an equivalent, purely mathematical claim M_q as follows:

$$M_q : \text{ ‘} f_x(0) < q, \text{ where } f_x \text{ is the } x\text{-coordinate of } f.\text{’}$$

We will then have that p_q is true iff M_q is true. It is a requirement of any mathematical representation of the trajectory of a particle that, for each p_q, there is some purely mathematical proposition M_q with this property.

Imagine that, using some sort of unusual mathematical machinery, we are able to construct a mathematical representation of the particle P. Using this machinery, we must be able to construct purely mathematical propositions M_q (presumably different from those constructed above) such that p_q is true iff M_q is true. (We do not place any restriction on the M_q – in particular, we do not even require that the M_q be first-order for what follows.) Now define the following subset S of the rationals $\mathbb{Q}^V$:

$$x \in S \text{ iff } x \in \mathbb{Q}^V \text{ and } M_x \text{ is true.}$$

Then S is just the Dedekind cut of $\mathbb{Q}^V$ corresponding to the real $r \notin \mathbb{R}^V$. So regardless of which type of mathematical structure we use to define the trajectory of P, the real $r \notin \mathbb{R}^V$ will be definable in turn from sentences concerning that mathematical structure.

What can we conclude from this? If S is any mathematical structure from which the trajectory of the point particle P can be defined, then S can be used to define a real – more specifically, a Dedekind cut of $\mathbb{Q}^V$ – not in $\mathbb{R}^V$. So if the trajectory of the particle P is mathematically representable, then the referent of $\mathbb{R}$ cannot be $\mathbb{R}^V$ – it must include some $r \notin \mathbb{R}^V$. To put it differently – if we have any mathematical machinery with which we can represent the trajectory of P, then we have the mathematical machinery with which we can show that there are reals that are not in $\mathbb{R}^V$. So regardless of what one takes the trans-world referent of $\mathbb{R}^V$ to be, one can construct a possible world containing a point particle P such that, if P can be mathematically represented in some way, then one can construct a real not in $\mathbb{R}^V$. It follows from the Mathematical Premise that the trajectory of such a particle *cannot* be represented by *any* mathematical structure. So the world w is *not* isomorphic to a giant mathematical structure, and the language of nature need not be mathematics. This concludes my main argument.

Having finished the main argument, it is worth contrasting it with a different family of arguments one finds in philosophy – namely, arguments that revolve around the phenomenon of 'unintended models'. Such arguments exploit the fact that any first order characterization of a structure like $\mathbb{N}$ or $\mathbb{R}$ will admit non-standard models. The fact that such structures cannot be *uniquely* characterized in a first-order way is then used to draw various conclusions of an ontological or epistemological nature. Perhaps the most famous example of this comes from Putnam – see pp 217-218 of [74].

The most significant difference between the argument of the present paper, and the arguments involving unintended models, is that while the later sorts of arguments revolve around the fact that consistent sets of sentences often have *multiple* models, my argument exploits the fact there are consistent sets of sentences that have *no* models. Of course, this requires us to go beyond standard first-order logic, as Godel's Completeness Theorem tells us that any consistent first-order set of sentences has a model. This is why we choose to work with sentences with *infinitary* connectives, because in such langauges, the analogue of Godel's Completeness Theorem is known to fail. For a discussion of this, see pp 135-136 and Appendix C of Dickman [26], and pp 333-334 of Scott [85]. Specifically, in the article by Scott, one finds reference to a construction of a complete, consistent set of sentences of $\mathcal{L}_{\omega_1,\omega}$ that has no model.

To make the point more explicit, let me construct the consistent set of sentences of $\mathcal{L}_{\omega_1,\omega}$ with no model, that is (in effect) used in the argument that I have just presented. (The set of sentences is different from those given in Dickman's [25] and Scott's [85].) Let $\{c_q\}_{q\in\mathbb{Q}}$ be a set of constants, where $\mathbb{Q}$ are the standard rationals. Let T_1 be the set of sentences consisting of $c_{q_1} + c_{q_2} = c_{q_3}$ (respectively, $c_{q_1}.c_{q_2} = c_{q_3}$) for all $q_1, q_2, q_3 \in \mathbb{Q}$ such that $q_1 + q_2 = q_3$ (respectively, $q_1.q_2 = q_3$). Let $i : \{c_q\}_{q\in\mathbb{Q}} \to \mathbb{Q}$ be the mapping defined by $i(c_q) = q$ for all $q \in \mathbb{Q}$. Let R be a binary relation, and let T_2 be the following set of sentences:

1. $\forall x, y[R(x, y) \to ((\bigwedge_{q\in\mathbb{Q}} x \neq c_q) \wedge (\bigvee_{q\in\mathbb{Q}} y = c_q))]$

2. $\forall x[(\bigwedge_{q\in\mathbb{Q}} x \neq c_q) \to \exists z R(x, z)]$

3. $\forall x[(\exists y R(x, y)) \to \exists z((\bigvee_{q\in\mathbb{Q}} z = c_q) \wedge \neg R(x, z))]$

4. $\forall x, y, z[(R(x, y) \wedge (z < y)) \to R(x, z)]$

5. $\forall x, y[R(x, y) \to (\exists z)((z > y) \wedge R(x, z))]$

The set of sentences T_2 says that, for any x not in $\{c_q\}_{q\in\mathbb{Q}}$, $\{y : R(x, y)\}$ is a Dedekind cut of $\{c_q\}_{q\in\mathbb{Q}}$, under its natural linear ordering. Now let d be a new constant symbol. For each $r \in \mathbb{R}^V$, let T_r be the following sentence:

$$(\bigvee_{z\in i^{-1}(r)} \neg R(d,z)) \vee (\bigvee_{z\notin i^{-1}(r)} R(d,z)).$$

Where $i : \{c_q\}_{q\in\mathbb{Q}} \to \mathbb{Q}$ is the isomorphism described above.

Consider the set of sentences S consisting of T_1, T_2, $\bigwedge_{q\in\mathbb{Q}}(d \neq c_q)$, and T_r for each $r \in \mathbb{R}^V$. Then we have the following:

Theorem: S is a consistent set of sentences that has no models in V.

Proof. That S is a consistent set of sentences follows from the Main Theorem. Assume now that S has a model in V, and let $r = i[\{y : R(d,y)\}]$. Then it is easy to see that r is a Dedekind cut of $\mathbb{Q}^V$ not contained in $\mathbb{R}^V$. Thus, $r \notin V$, which is a contradiction. □

The set of sentences S is essentially the one exploited in my main argument. Note also that the set S of sentences is not only consistent, but consistent with ZFC, unlike the consistent sets of sentences with no models referred to in the works of Dickman and Scott.

Having completed my argument, in the next two sections (§§3–4) I shall discuss some objections. In §3 I shall focus on objections that concede the Modal and Mathematical Premises, but that argue that my conclusion somehow fails to follow. In §4 I will then discuss the obstacles involved in denying either the Mathematical or Modal premise.

6.3 OBJECTIONS AND REPLIES.

In this section, I wish to consider several objections to the validity of the argument just given. The first objection really just amounts to a suspicion that something has gone wrong:

Objection 1: You appear to have used mathematics to construct something mathematically ineffable. There is something paradoxical about this - something must have gone wrong somewhere in your argument.

Reply: Contrary to appearances, no such thing has been done. The main mathematical step in my argument is simply directed towards showing:

$$Con(ZFC) \rightarrow Con(ZFC + \text{`there is a Dedekind cut of } \mathbb{Q}^V \text{ not contained in } \mathbb{R}^V.\text{'})$$

This is, first and foremost, a claim about sentences, and does not involve the construction of anything mathematically ineffable in any way. For instance, this mathematical result does not involve the explicit construction of a model of ZFC satisfying 'there is a Dedekind cut of $\mathbb{Q}^V$ not contained in $\mathbb{R}^V$.'

The real 'constructive' work in the argument of the previous section is done by the Modal Premise, according to which every logically consistent state of affairs is possible. From this, we are able to move from a consistent set of sentences to a possible world in which those sentences are true. This is not a 'mathematical construction' in any sense of the word – if anything, it is a 'philosophical construction'. It is from this 'philosophical construction' that something mathematically ineffable can be found. The mathematically ineffable trajectory has therefore been 'philosophically', and not 'mathematically', constructed.

Objection 2: Your techniques prove too much. In particular, they show that mathematical truths cannot be necessary truths. To see why, let w_1 be a possible world in which space-time has the Euclidean structure $(\mathbb{R}^V)^4$. Using the Main Theorem together with the Modal Premise, construct a possible world w_2 in which space-time has the Euclidean structure $(\mathbb{R}^*)^4$, where $\mathbb{R}^V \subset \mathbb{R}^*$, and in which $\mathbb{R}^*$ contains a real number not contained in $\mathbb{R}^V$. Consider the proposition Y defined as follows:

$$\text{Y}: \text{ `}\exists x \in \mathbb{R}\ [x \notin \mathbb{R}^V]\text{'}$$

Then in w_1, Y is false, but in w_2, Y is true. Mathematical truths are no longer necessary truths.

Reply: One option here is to go along with the objection, and conclude that mathematical truths are not necessary truths. I will not take this route for two reasons: first – I think that there are independent reasons for wanting mathematical truths to be necessary, and second – we shall see in the next section that someone who does not believe in the necessity of mathematical truths is unlikely to go along with the Mathematical Premise. We must consequently face this objection head on.

I do *not* think that the techniques of the previous section show that mathematical truths are not necessary truths. Let us consider the possible world w_2 described in the objection. If one held that the 4-dimensional Euclidean geometry of space time w_2 *had* to be mathematically representable (i.e., had to be isomorphic to some fixed mathematical structure), then it could indeed be inferred that such a structure would have to have the form $(\mathbb{R}^*)^4$, where $\mathbb{R} \subset \mathbb{R}^*$. This would entail the existence of real numbers not in $\mathbb{R}^V$. One could then take this to mean that the sentence Y was true in w_2 and false in w_1.

However, this reasoning is based on the assumption that the 4-dimensional Euclidean geometry of space-time has to be fully mathematically representable. This is precisely the view I want to reject. If one does *not* insist that the 4-dimensional Euclidean geometry of space-time must be mathematically representable, then one can say that, in w_2, there is a portion of space-time (specifically, the portion not contained in $(\mathbb{R}^V)^4$) that is mathematically ineffable. This does not mean that the sentence Y is true in w_2 – to the contrary, insofar as one focuses on the portion of space time that *does* admit a mathematical description, Y will be false. That is, insofar as in the sentence

$$\text{Y}: \text{ `}\exists x \in \mathbb{R}\ [x \notin \mathbb{R}^V]\text{'}$$

one takes the existential quantifier to range over the mathematically describable points of space time, Y is false. Of course, one is free to let the existential quantifier range over the 'points of space-time in general', in such a way that the mathematically ineffable points are included; but in that case, Y is no longer a purely *mathematical* assertion – it is a hybrid mathematical/ ordinary language assertion – and therefore no tension is to be found with the doctrine that the truths of mathematics must be necessary.

In brief, it is only insofar as one clings to the idea that all physical structures must be mathematically representable that one can reach the uncomfortable conclusions outlined in the objection. Once one rejects this idea, as I wish to do, the uncomfortable conclusions outlined in the objection are no longer forced upon us.

Objection 3: Your argument exploits the fact that ZFC is unable to fix the reference of $\mathbb{R}$. However, perhaps the trouble here lies with ZFC. If you want to insist that the reference of $\mathbb{R}$ is fixed, perhaps we need to add axioms to ZFC to guarantee this – for instance, perhaps we should consider adding an axiom that specifically tells us what the referent of the term $\mathbb{R}$ is to be. Instead of pointing out an expressive limitation of mathematics itself, your argument does nothing more than highlight a problem with ZFC.

Reply: The problem with this suggestion is that there is *no* way of extending ZFC to a larger first-order theory ZFC^* in such a way that ZFC^* fixes the reference of $\mathbb{R}$. To see this, let c be a constant symbol unused in ZFC^*, and let T be the set of sentences consisting of ZFC^*, '$c \in \mathbb{R}$', and '$c \neq r$' for each $r \in \mathbb{R}$. Then the set of sentences T is consistent, because each finite subset of T is consistent. So any model of ZFC^* will contain a real number not in $\mathbb{R}$. Consequently, ZFC^* does not fix the reference of $\mathbb{R}$. The inability to fix the referent of $\mathbb{R}$ is not a peculiarity of ZFC, but applies to any first-order set theory quite generally. New axioms cannot save us.

In reply to this, it could be suggested that a more radical departure from conventional wisdom in the foundations of mathematics is needed. For instance, perhaps the foundations of mathematics need to be stated in a higher order logic, or infinitary logic, for which we do not have the usual compactness, completeness, and incompleteness results on which the previous arguments depend. On one level, of course, I have nothing with which to defend myself against this specific line of attack – if my opponent wants to radically rewrite the foundations of mathematics, then I have no desire to stop him.

Nevertheless, it is worth noting that there may be good reasons for thinking of first order logic as the appropriate logic in which to seek the foundations of mathematics. Consider, for instance, second-order logic, in which quantifiers of the form $\forall X \subseteq Y$ and $\exists X \subseteq Y$ are taken as basic. One might worry that it is precisely the job of the foundations of mathematics to explain what the relationship '$\subseteq$' consists in, rather than taking it as unanalyzable. One might express similar qualms about infinitary connectives. Such concerns might lead one to think that only first order logic can provide the logical framework for the foundations of mathematics, and that higher order logics cannot fulfill this role.[8] If this is right, then there is perhaps less room than one might think for arguing that my result highlights nothing more than an inadequacy in the current foundations of mathematics.

6.4 MORE OBJECTIONS AND REPLIES.

In this section, I wish to consider whether or not there are simple ways to deny either the Modal or Mathematical Premises, and therefore avoid the conclusion of my argument.

Let us begin by considering the Modal Premise, which states that every logically consistent state of affairs is possible. There are various ways in which one could weaken this claim in order to avoid the conclusion of my main argument. For instance, one could maintain that, in full generality, the Modal Premise is false, and should be replaced by the following claim:

Weak Modal Premise: A logically consistent set of sentences is possible just in case it describes a possibility that can be represented mathematically.

One would therefore be able to avoid the conclusion that there might be mathematically ineffable structures in the world. Thus, we have:

Objection 4: By endorsing only the Weak Modal Premise, the conclusion of the main argument can be avoided.

Reply: My problem with the Weak Modal Premise is that it places an artificial restriction on which states of affairs are to count as possible.

Consider an analogy. Imagine someone who, in order to convince us that it is impossible for evil to triumph over good, tells us that a logically consistent state of affairs is possible just in case it involves evil triumphing over good. If he cannot justify why, on *independent grounds*, we should accept his restriction on the nature of possibility, then he cannot hope to be taken seriously. How does advancing the Weak Modal Premise differ from this?

[8] Whether this conclusion is nothing other than the expression of a philosophical and mathematical prejudice is an open question that I cannot hope to address here.

One might suggest that there *are* independent grounds for thinking that we ought to restrict the Modal Premise to cover only situations that are mathematically representable, and deny the possibility of situations that cannot be mathematically represented. In the first section of this chapter, I presented one such argument, and showed it to be unpersuasive. As an alternative, one might suggest that mathematics *is* the science of all possible structures, and therefore the restriction implicit in the Weak Modal Premise is reasonable. But I do not think that this is persuasive – if mathematics really *is* the science of all possible structures, then one should not *need* to add provisos to the Modal Premise in order to block the existence of mathematically ineffable structures – one should be able to get any relevant provisos, and block the existence of mathematically ineffable structures, for 'free'. What my argument shows is that one *cannot* blindly assume that mathematics is the science of all possible structures. To try and salvage the hypothesis that mathematics is the science of all possible structures by restricting the world of possible structures to those which we already know are mathematically representable, is no different from the awkward maneuvers described in the above paragraph concerning the triumphing of good over evil.

I therefore think that it would be unreasonable to try and block the possible existence of mathematically ineffable structures by moving from the Modal Premise to the Weak Modal Premise. Such a move simply seems ad hoc. Of course, to say that something is ad hoc is not to argue that it is false. There is no *logical* problem with the Weak Modal Premise. However, the arbitrariness of the restrictions involved make the Weak Modal Premise an unattractive alternative to the Modal Premise.

Let us consider another objection to the Modal Premise.

Objection 5: There are all sorts of logically consistent statements that do not describe possibilities. For instance, (i) that there are natural numbers unequal to $1, 2, 3, \ldots$ (where the list just given consists of the standard naturals) or (ii) that there are reals not in $\mathbb{R}^V$, are logically consistent statements that do not describe bona fide possibilities. (This point is made by Almog [1], in a discussion of a different matter.) But surely the main argument in question exploits precisely such possibilities – in particular, (ii). The main argument therefore uses the Modal Premise in precisely the sort of situation where we have reason to think it fails.

Reply: The Modal Premise, as I have used it, states that every logically consistent state of affairs is possible. However, I do not think that facts about logic and mathematics are 'states of affairs' in this sense – they are best thought of as truths that govern *all* possible worlds, rather than facts that may or may not hold in *particular* possible worlds. Thus, I would not consider the law of non-contradiction, or the fact that $2 \times 2 = 4$ to be 'states of affairs'. The proposition X to which I apply the Modal Premise can, in contrast, be spelled out in purely operationalistic terms, and thus is the sort of thing to which it is entirely fair to apply the Modal Premise.

Moreover, there is no sense in which my argument assumes (or entails) that it is possible that there are reals not in $\mathbb{R}^V$. This is true only if one assumes that the 4-dimensional Euclidean geometry of space time has to be mathematically representable (in particular, as the set of Dedekind cuts of some ordered field $\mathbb{F}$.) But the mathematical representability of space-time (in particular, of the space-time in the possible world in which X is true) is precisely what I wish to challenge. Thus, for reasons similar to those given in reply to Objection 2., there is no real sense in which my argument entails the existence of reals not in $\mathbb{R}^V$.

What about the Mathematical Premise? One might think that we are at greater liberty to deny this assumption. I think, however, that the obstacles surrounding the denial of the Mathematical Premise are actually more substantial that those surrounding the denial of the Modal Premise.

The main problem with denying the Mathematical Premise is that it forces us to give up the idea that mathematical truth is a species of necessary truth. To see why, assume that the Mathematical Premise is false. Let w_1 and w_2 be possible worlds in which $\mathbb{R}^{w_1} \neq \mathbb{R}^{w_2}$, where $\mathbb{R}^{w_1}$ and $\mathbb{R}^{w_1}$ are the referents of the term 'the reals' in w_1 and w_2. Assume, without loss of generality, that $\exists x[x \in \mathbb{R}^{w_2}$ and $x \notin \mathbb{R}^{w_1}]$.

Consider the proposition Z defined as follows:

$$\text{Z}: \text{ `}\exists x \in \mathbb{R}\ [x \notin \mathbb{R}^{w_1}]\text{'}$$

Then in w_1, Z is false, but in w_2, Z is true. So mathematical truths are no longer necessary truths.

One might think it unfair that, in our sentence Z, we have made explicit reference to a possible world w_1. I do not think this is anything to be worried about, as we can just think of $\mathbb{R}^{w_1}$ as the name of a set, rather than a definite description involving the non-mathematical term w_1. However, it is worth discussing ways in which this difficulty may be able to be avoided.

Using the method of set theoretic forcing, one can 'add reals' in such a way that the value of the continuum is changed. (For references on the method of forcing, see [48] or [53].) Specifically, we begin with the following theorems (proofs for which can be found in the references just given):

Theorem 1. $Con(ZFC) \rightarrow Con(ZFC + 2^{\aleph_0} = \aleph_1)$.

Theorem 2. $Con(ZFC) \rightarrow Con(ZFC + 2^{\aleph_0} = \aleph_2)$.

What these theorems show is that it is possible for the definite description $\mathbb{R}$ to refer to sets $\mathbb{R}^*$ and $\mathbb{R}^{**}$ respectively, such that the following holds:

(i) the set $\mathbb{R}^*$ has no uncountable subset S such that $card(S) \neq card(\mathbb{R}^*)$, and

(ii) the set $\mathbb{R}^{**}$ has an uncountable subset S such that $card(S) \neq card(\mathbb{R}^{**})$.

Consider now the proposition Y:

$$\mathrm{Y}: \text{ `}\exists S \subset \mathbb{R}\ [S \text{ is uncountable, and } card(S) \neq card(\mathbb{R})]\text{'}$$

If $\mathbb{R}$ refers to $\mathbb{R}^*$, then Y is false, but if $\mathbb{R}$ refers to $\mathbb{R}^{**}$, then Y is true. This is an example of a mathematical assertion not explicitly or implicitly making reference to any possible world, which can perhaps be contingently true, but not necessarily true.

Of course, it is perfectly consistent to maintain that, although the term $\mathbb{R}$ can have different references in different possible worlds, it cannot refer to both $\mathbb{R}^*$ in one possible world, and $\mathbb{R}^{**}$ in another. (More generally, in order to avoid the sort of problem I am trying to make here, one would have to insist that the various possible references of the term $\mathbb{R}$ all be 'elementarily equivalent'; see [19] for an explanation of this terminology.)

But in that case my first counterexample to the necessity of mathematical truth given by the proposition Z still holds. In order to deny the Mathematical Premise, one therefore must either (i) accept that there can be mathematical truths that are not necessary, or (ii) argue that the various references of the term $\mathbb{R}$ are all elementarily equivalent, *and* that there is something illegitimate about the mathematical proposition Z. Both of these options are possible; the first, however, parts substantially with conventional philosophical wisdom about mathematical truth, while the second cries out for a detailed technical account, yet to be provided, of the possible references of mathematical terms. Neither of these options are absurd, but the pursuit of either option demands, I think, substantial developments in the philosophy of mathematics. Neither option offers an easy and obvious 'safe haven' from my main argument.

Finally, I wish to consider the following explicit objection to the Mathematical Premise:

Objection 6: The Mathematical Premise should be rejected, because your argument tells us that it should be rejected. In particular, let us assume that $\mathbb{R}$ is the unique transworld referent of the term 'the reals'. With the use of the Modal Premise, consider the possible world w in which the proposition X (defined in Section 2.2) is true. Let $S = \{q : q \in \mathbb{Q}$ and the particle is to the left of $x = q$ at $t = 0\}$. Then S defines a real not in $\mathbb{R}$, and so the referent of 'the reals' in w cannot be $\mathbb{R}$. The Mathematical Premise is therefore false.

Reply: Much like my replies to previous objections (in particular, objections 2 and 5), the confusion implicit in this objection arises from supposing that the structure of space in w corresponds to some mathematical structure. But such a supposition is precisely what I wish to challenge.

In particular, we cannot take for granted that the predicate:

'is a rational number q such that the particle in w is to the left of $x = q$ at $t = 0$'

defines a set S. One might think that such a set S *must* exist, for the following reason: define $R(q)$ iff q is a rational number such that the particle in w is to the left of $x = q$ at $t = 0$'. Now define, for each natural number j,

$$S_j = \{q_i : i \leq j \text{ and } R(q_i)\},$$

where q_i is some fixed enumeration of the rationals. Each S_j is a finite set that can be physically written down, and thus exists. But $S = \cup_j S_j$, and so S also exists. Thus, from the fact that if a sequence of sets exists, its union exists, we can conclude that S exists.

The problem, however, is that we have not argued that the *sequence* $< S_j >$ exists – all we have argued is that each S_j *individually* exists. This is an important distinction – one can construct models V, V' of ZFC (assuming $Con(ZFC)$) such that $V \subset V'$, and such that there exists sets $S_i \in V$ (for i an arbitrary natural number) such that $\cup_i S_i \notin V$, and $\cup_i S_i \in V'$. Unless one can argue for the existence of the sequence $<S_j>$, one cannot argue for the existence of S.

For another attempt to mathematically define S, one might argue that the existence of S follows from the Axiom (Schema) of Separation, according to which, for any predicate ϕ in the language of set theory, for any set X, and any set p, the set $\{u \in X : \phi(u, p)\}$ may be taken to exist. In this case, we would simply let X be the set of rationals, and define $\phi(u)$ iff $R(u)$ (ignoring the parameter p.)

The problem, however, is that R is not a predicate in the language of set theory, for at least two reasons. First, it involves reference to the term w, which is not a purely mathematical term. But second (and more importantly), it involves the ordinary language term 'to the left of'. One might think that one can formalize 'to the left of' by associating each point of space with some element of a set A, and defining some relation $<_A$ on A. But to assume this is simply to beg the question – i.e., to assume this is to assume that the structure of space in w *can* be mathematically represented. But this is presicely the sort of assumption that my main argument challenges.

As far as I can see, all attempts to mathematically define such an S, and thereby perform a reductio against the Mathematical Premsise, presuppose that the language of nature *must* be mathematics. The argument given in the Objection is therefore unpersuasive.

Let us consider one final objection before moving on.

Objection 7: The argument I have offered presupposes that we may erect a plane at $x = q$ for infinitely many values of q. It is from this infinite sequence of measurements that the problematic 'Dedekind cut' arises. However, in physics, when conceptual problems of this sort arise as a consequence of an assumption, we reject the assumption – for instance, because of the problem in talking about the energy of point particles in classical electrodynamics, it is generally declared that, at least when discussing point particles, the energy concept should be abandoned. In the present case, we should do the same thing, and conclude that the concept of 'infinitely many measurements' (or something like that) should be abandoned.

Reply: First of all, when constructing our one particle possible world that is not isomorphic to any mathematical structure, we are free to stipulate (i) that the particle remains fixed in its location throughout time, and (ii) that accurate measurements are performed on the particle at $t = 0, 1, 2, ...$ for all possible different values of q. (If we take this second clause seriously, then our possible world is no longer really a one particle world, because it now has a little man running around doing measurements, but this does not matter.) The applicability of the concept of 'infinitely many measurements' is then built in to the world in question, and so cannot be dismissed as somehow inappropriate.

But second of all, even if we allow the particle to wander off after $t = 0$, and we only perform one measurement at $t = 0$, we are always free to talk about the result that would have occurred were we to have performed a different measurement at $t = 0$ – i.e., we are always free to talk about the results of *counterfactual* experiments at $t = 0$, in spite of the fact that only one measurement actually occurs at that time. This sort of move seems to be an indispensable part of physics – for instance, in spelling out the physical content of the assertion that the wavefunction of a given system is ϕ at a given time, reference will have to be made to not only the outcome of some actual measurement performed on the system at that time, but also to the results of counterfactual experiments performed on the system at that time.

The important point is that we do not need the concept of 'infinitely many measurements actually being performed' to get the argument to work – all we need is the concept of a *counterfactual measurement*, which is so fundamental to all physical theories, that to dismiss it in order to avoid the conclusions of this chapter would be to throw the baby out with the bathwater.

6.5 A COROLLARY.

Before discussing the connection between my main argument and the Modest Deduction Failure Thesis, it will be helpful to discuss an interpretative error that could arise at this point.

In the previous sections, I have argued that the trans-temporal universe need not be isomorphic to any mathematical structure. One might interpret this result by first suggesting that the set of sentences T true of the trans-temporal, physical world need not have a model, and then going on to suggest that the use of mathematics in physics may well have to be restricted to a purely syntactic, rather than semantic role. That is to say, although we can certainly use the *language* of mathematics to *list* the propositions true of the physical world, we cannot necessarily use the *structures* of mathematics to *model* the physical world.

However, this way of interpreting my main result is not correct. Specifically, it runs afoul of Godel's completeness theorem, (see [19]), according to which every consistent set of first-order sentences has a model. I will spend this section discussing a subtlety surrounding this point. My conclusion will be that, if the results of the previous section are accepted, then the language of mathematics cannot necessarily be used to even *list* the first-order propositions true of the physical world. It is not just the mathematician's ability to appropriately *model* some set of first-order sentences that is being challenged; rather, the ability of the mathematician to *form* such a set of first-order sentences in the first place is equally under attack.[9]

Godel's completeness theorem tells us that if we can construct a consistent, complete[10] set of first-order sentences T true of physical reality, then we can also construct a model M for T; that is to say, we can construct a mathematical structure M in which the sentences of T are true. If it is not necessarily true that the trans-temporal universe is isomorphic to any mathematical structure, then it follows that it is not necessarily true that there exists a consistent, complete set of sentences T true of physical reality.

One might, however, have the following concern about the argument just given: the

[9]It follows, of course, that we cannot necessarily form the set of sentences with infinitary conjunctions/ disjunctions true of trans-temporal, physical reality either.

[10]A set of sentences T is complete just in case for any sentence s, either $s \in T$ or $\neg s \in T$.

model of T (i.e., the model of the set of first-order sentences true of physical reality) that we construct with the help of Godel's completeness theorem need not be isomorphic to physical reality. Thus, from the fact that the set of first-order sentences T true of physical reality is a bona-fide set, it does not follow that the world is isomorphic to a mathematical structure. Generally speaking, T will have all sorts of models, many of which will contain elements with no physical interpretation. How do we know that T has *any* model with the property that each element of physical reality has an interpretation in the model, and that each element of the model corresponds to some actual element of physical reality? If the existence of such a model of T cannot be guaranteed, then we *cannot* use Godel's completeness theorem to argue that once we have constructed a consistent, complete set of first-order sentences T true of physical reality, then we can also construct a mathematical structure to which the trans-temporal universe is isomorphic. This would then still leave room for the interpretation of my main result offered in the first paragraph of the present section.

Given a reasonable assumption about T, I claim however that we *can* guarantee that T has a model in which every element of physical reality has an interpretation, and in which every element of the model corresponds to some actual element of physical reality. To demonstrate this, I shall first need a definition (taken from [19]):

> Let T be a set of first-order sentences, and C a set of constant symbols. Say that C is a *set of witnesses* for T iff for every first-order formula ϕ with at most one free variable, there is a constant symbol $c \in C$ such that $\exists x \phi(x) \to \phi(c)$ is a logical consequence of T.

Now, let N be the set of names n_x for all elements of physical reality x,[11] and let T be the set of sentences true of physical reality. We assume that N is a set of witnesses for T – that is to say, we assume that if some existential claim is true of physical reality, then we can name some element of physical reality which witnesses that the existential sentence in question is true. We take the claim that N is a set of witnesses for T to be a requirement of any complete description of physical reality.

Consider the following result (Lemma 2.1.2 from [19]):

> Let T be a consistent set of sentences and C be a set of witnesses for T in a language L. Then T has a model such that every element of that model is the interpretation of a constant $c \in C$.

[11] If, for instance, points of space-time are to be regarded as real, then we must assume that they are given names too.

From this, the following may be deduced: if T is a complete set of sentences describing physical reality, N is a set of names for all the elements of physical reality, and N is a set of witnesses for T, then it is possible to construct a model of T such that each element of physical reality has an interpretation in the model, and each element of the model corresponds to some actual element of physical reality.

Thus, from a complete, consistent list of the sentences true of physical reality, it is possible to build a mathematical structure isomorphic to physical reality. As a corollary to my main result, we see that it must not necessarily be possible to even *form* the set of first-order sentences true of trans-temporal, physical reality. The interpretation of my main result offered at the beginning of the present section is therefore incorrect.

6.6 THE REPRESENTATION FAILURE THESIS.

Physics tries to make sense of the world using mathematics. In fact, the trend of post-Aristotelian physics strongly suggests that all there is to be known about the world can be expressed in purely mathematical language. What implications do the arguments presented thus far have for this doctrine? Consider the following:

Theorem: There is no function $\phi : \mathbb{Q} \to \{0, 1\}$ such that $\phi(q) = 1$ if, at $t = 0$, the particle in the world w (described in section 2.2) lies to the left of a plane erected at $x = q$, and $\phi(q) = 0$ otherwise.

Proof. If such a ϕ existed, we could use the set $\phi^{-1}(1) = \{q : \phi(q) = 1\} \in V$ to define a Dedekind cut of $\mathbb{Q}$ both in V and not in V, contradiction. □

If there were some body of knowledge that could be used to completely describe the position of the particle in world w over time, then there would have to be a function ϕ : $\mathbb{Q} \to \{0, 1\}$ such that $\phi(q) = 1$ if, at $t = 0$, the particle lies to the left of a plane erected at $x = q$, and $\phi(q) = 0$ otherwise. The Theorem just cited, however, shows that such a function ϕ does not exist. Thus, there is no complete mathematical theory of the world w.

Recall the Representation Failure Thesis:

The Representation Failure Thesis: We cannot take for granted that the world is representable in mathematical terms – that is to say, we cannot take for granted that the state of the world, together with all its fundamental laws, can be captured perfectly by some set of mathematical assertions or equations. In fact, it is more likely that the world is *not* representable in mathematical terms (in this sense) than that it is.

I have argued that we cannot take for granted that the world is representable in mathematical terms (in the sense described.) But how big is the set of possible worlds for which the language of nature is mathematics? If a possible world is selected at random, is it likely that the language of nature is mathematics in such a world? Unlikely? Neither? Until we have addressed this question, our defense of the Representation Failure Thesis is incomplete. In the remainder of this chapter, I shall argue that it is far more likely that the world is *not* representable in mathematical terms than that it is, thereby completing my argument for the Representation Failure Thesis.

Say that a possible world w is a one-particle, Euclidean world just in case:

1. The space-time of w has the geometry $\bar{\mathbb{R}}^4$, where $\bar{\mathbb{R}}$ is an ordered field consisting of a set of Dedekind cuts of $\mathbb{Q}^V$.

2. The world w contains exactly one particle, and this one particle exists throughout all time in w.

Some one-particle, Euclidean worlds will be isomorphic to structures in V, while others (such as the one constructed in my main argument) will not. We have shown how to construct one-particle, Euclidean worlds not isomorphic to any structure in V with the use of forcing. One can 'iterate' the forcing construction to obtain further worlds not isomorphic to any structure in V. Specifically, beginning with V and $\mathbb{R}$, we may iterate the forcing construction, adding a real at each ordinal level α, obtaining a sequence $\mathbb{R} = \mathbb{R}_0 \subset \mathbb{R}_1 \subset ... \subset \mathbb{R}_\alpha \subset ...$ of possible referents for the reals. (For details on the technique of iterated forcing, see Section 23 of Jech [48] or Chapter VIII of Kunen [53].) Each $\mathbb{R}_\alpha$ can then be used to construct a world w_α such that the geometry of space-time in w_α is given by $\mathbb{R}_\alpha^4$, and each w_α is distinct. Thus, for each α, we have a (distinct) possible world w_α not isomorphic to any structure in V. From this, we may conclude that the set of one-particle, Euclidean possible worlds not isomorphic to any structure in V (i.e., the set of one-particle, Euclidean worlds for which the language of nature is *not* mathematics) forms a proper class.

On the other hand, the set of possible one-particle, Euclidean worlds that *are* isomorphic to a structure in V forms a bona-fide set (this is because the requirement that $\bar{\mathbb{R}}$ is an ordered field consisting of a set of Dedekind cuts of $\mathbb{Q}^V$ guarantees that the cardinality of $\bar{\mathbb{R}}$ is less than or equal to the continuum.) Thus, the set of one-particle, Euclidean worlds *not* isomorphic to a structure in V vastly outnumbers the set of one-particle, Euclidean worlds that *are* isomorphic to a structure in V. In this particular sense, then, we have an answer to the question posed above – if a one-particle, Euclidean world is selected at random, it would be much wiser to bet that it was *not* isomorphic to a structure in V than that it was.

So, given the information that the world w is a one-particle Euclidean world, and no other knowledge, we have good (although not decisive) reason to expect that w will not be isomorphic to a structure in V.

Of course, the situation in reality is very different. We know that the world is not a one-particle world, and we also have good reason to think that it is not Euclidean (in the sense defined). But consider the following argument – fix a cardinal κ, and let S_κ be the set of mathematical structures in V which qualify as candidates for models of the actual world, given everything known about the actual world. Let us assume for now that space-time is homeomorphic to some complete, separable space with no isolated points – these are known as *Polish* spaces, and their set theoretic properties have been thoroughly investigated; see Moschovakis [63], for example. Such a Polish space is determined (up to isomorphism in V) by a separable subset S, together with the metric on S. For a cardinal κ, let P_κ be the set of possible worlds in V which:

1. are consistent with everything we know about the world, and

2. have space-time homeomorphic to a Polish space which is determined by some S in V_κ.

Then P_κ forms a bona-fide set. However, consider the collection C_κ of possible worlds w *not* isomorphic to a structure in V which:

1. are consistent with everything we know about the world, and

2. have space-time homeomorphic to a Polish space with a separable subset S in V_κ.

Then C_κ is a class, because we can add points to such a Polish space by the method of forcing, and iterate this construction through the ordinals, as before.

Thus, in this sense, the set of possible worlds that (i) are consistent with everything we know, and (ii) are *not* isomorphic to a structure in V, vastly outnumbers the set of possible worlds that are (i) consistent with everything we know, and (ii) *are* isomorphic to a structure in V. In this particular sense, then, if a world consistent with everything we know was selected at random, it would be much wiser to bet that it was *not* isomorphic to a structure in V than that it was. Thus, again, we have good (although not decisive) reason to expect that the actual world is not isomorphic to a structure in V. The Representation Failure Thesis follows from this.

7.0 UNREASONABLE EFFECTIVENESS: HISTORICAL/AESTHETIC.

7.1 INTRODUCTION

In this chapter, I would like to begin my discussion of the problem of the unreasonable effectiveness of mathematics in physics. This discussion will occupy both chapters 7 and 8. In this chapter, I will discuss several aspects of the problem that do *not* specifically relate to the Representation Failure Thesis, saving my discussion of the aspects of the problem that *are* connected with the Representation Failure Thesis for the next chapter.

In [101], the physicist Eugene Wigner argued that there is something surprising about the way in which mathematics is so useful in physics. Finding himself unable to explain why this should be so, Wigner declared the usefulness of mathematics in physics both a 'mystery' and a 'miracle'.

Unsurprisingly, this claim produced a flood of responses by scientists (e.g. [41], [98]) and mathematicians (e.g. [56], [94]) of the highest caliber. Some argued that the usefulness of mathematics in articulating the laws of nature was not really surprising at all. Some (e.g. [28], [105]) challenged the idea that mathematics *was* so extraordinarily useful to the physicist.

The philosophers refused to be left out of the debate. Some (e.g. [5], [35], [64]) argued that the usefulness of mathematics was not surprising, some (e.g, [91, 92]) argued that it was, and others (e.g., [103]) challenged the idea that mathematics was as useful as everyone else had said. The debate amongst philosophers closely paralleled the debate in the scientific and mathematical communities.

In order to get a grip on this debate, one might first try to divide the warring factions into three groups: those who think there is a mystery but cannot solve it, those who think there is a mystery and claim to be able to solve it, and those who think there is no mystery at all. A closer look at the literature, however, shows that this may not be the most useful way to carve things up. The problem is that most authors have taken an idiosyncratic view of the way in which mathematics might be regarded as unreasonably effective, and accordingly defended or attacked what they took the mystery to be. The debate has consequently involved a great deal of discussion at cross purposes. The tripartite division suggested above therefore turns out to be not especially useful; most people, after all, think that there is *some* sense of the mystery which is illegitimate, and other senses in which it is legitimate – even if solvable.

I therefore propose to distinguish several different ways in which mathematics could be taken to be 'unreasonably effective' in physics. I shall argue that there are at least three different ways in which the role mathematics plays in physics might be taken to be unreasonably effective. I shall then assess each of these three ways individually – the first two in this chapter, and the third in the next chapter.

First, one might think it mysterious that in many cases, concepts developed by mathematicians with little regard for physics have turned out to be indispensable in articulating the fundamental laws of physics. Indeed, it is said that the physicist often finds the most appropriate tools for his work amongst the mathematical concepts already developed and explored by mathematicians, many of whom had no eye for the needs of physics. I shall call this the *historical* problem of the unreasonable effectiveness of mathematics in physics.

On the other hand, some have thought that it is mysterious that the *values* of the mathematician – the mathematician's conception of what is elegant, natural, convenient, and so on – should be so useful to the physicist. These values, it is often said, have played a great role in scientific progress – especially in physics. But these values are also often said to be aesthetic (or even anthropomorphic) in nature, and therefore are not the sort of thing we should really expect to impact the progress of physics. Yet it seems that they have, and continue to do so. I shall call this the *aesthetic* problem of the unreasonable effectiveness of mathematics in physics.

Finally, some have thought that the fact that the fundamental laws of nature are expressible in the language of mathematics at all is itself a source of mystery. This final mystery I shall call the *descriptive* problem of the unreasonable effectiveness of mathematics in physics.

The main job of this and the next section will be to assess the merits of each of these 'mysteries' individually, and determine whether they are answerable riddles, pseudo-problems, or genuine unsolved mysteries. We will see that different versions of the problem come out differently on this score.[1]

In the next section of the paper, I will examine the historical problem, and argue that it is a pseudo-problem. Specifically, I will argue that the historical problem is based on historically inaccurate storytelling about the relationship between mathematics and physics. In the third section, I will examine the aesthetic problem, and argue that it too is a pseudo-problem resting on a misrepresentation of the methodology of physics. I then conclude the chapter with a discussion of some other candidate mysteries concerning the unreasonable effectiveness of mathematics in physics, and argue that they do not warrant our attention. The discussion of the descriptive problem then occurs in Chapter 8.

7.2 THE HISTORICAL PROBLEM.

7.2.1 Introduction.

In this section, I shall discuss the historical version of Wigner's problem.

It would be slightly odd if a mathematician, searching for intellectual distractions, ignorant of the needs of the physicist, and guided by nothing other than his own caprice, came up with a mathematical structure or theory that was later recognized to be part of the deepest structure of physical nature. A single occurrence of this sort would be a happy co-incidence, but not obviously in need of explanation.

It would be somewhat more perplexing, however, if at *many* times in the history of physics, when physicists found themselves struggling to articulate new laws, they were able to find precisely the structures and theories they needed already warmed over in the mathematician's bag of playthings. This repeated co-incidence would surely call for further explanation.

[1]It is difficult to say which one of these three mysteries Wigner was really concerned with. One conjecture is that there is no fact of the matter about which mystery Wigner intended, and that he vacillated between different versions of a problem which he had not precisely formulated.

In fact, precisely this seems to have been the case. See page 451 of Steiner [91]:

> *"Appolonius's theory of conic sections had no applications in physics until Kepler used it, fifteen hundred years later, in formulating the laws of planetary motion. Riemann developed non-Euclidean geometry to solve problems in thermodynamics; fifty years later, Einstein applied it to gravitation. Elie Cartan, in his doctoral dissertation, classified the simple Lie algebras not realizing that he was, at the same time, classifying elementary particles. The list is very long."*

In addition, see Weinberg [98] page 725:

> *"The mathematical structures that arise in the laws of nature, as far as we know them, at the deepest level that we know them, are often mathematical structures that were provided for us by mathematicians long before any thought of physical application arose. It is positively spooky how the physicists find the mathematician has been there before him or her."* [2]

Of course, it is not as if the physicist is restricted to the mathematician's bag of playthings – if he were, the above claims would offer no surprise at all. The physicist is entirely free to pursue mathematical avenues previously shunned by the mathematician. Yet remarkably, this is often unnecessary. The quantum theorist, for instance, finds Hilbert space theory already in a mature enough form for his use – almost as if it had been prepared in advance for him. The particle physicist's experience with group theory is much the same.

This problem – the amazing way in which the physicist finds that the mathematician has been there before him – was one of the problems that struck Wigner in [101]. In order to distinguish this problem from other similar problems, I shall call it the *historical problem of the unreasonable effectiveness of mathematics* – or just the *historical problem.*

In this section, I will argue that the sense of wonder often associated with the historical problem rests on an inaccurate reading of the history of physics. The idea that the physicist, browsing casually through the journals and textbooks of the mathematician, has been able to find exactly what he has needed to solve his latest physics problems, is a romanticized misrepresentation of discovery in physics. Although there is much trafficking of ideas between mathematicians and physicists, we must be careful not to paint too rosy a picture of the ease with which this dialogue occurs. Even in the best cases, the physicist must struggle tooth and nail to make existent mathematics relevant to his work, as we shall see.

---

[2] In the same article, Weinberg [98] page 727 draws the following apt comparison:

> *"It is as if Neil Armstrong, when he first arrived on the moon and stepped out of the Apollo landing module, found in the lunar dust in front of him the footsteps of Jules Verne. "*

In this section, I wish to make three claims about the relationship between mathematics and physics, each of which challenges the idea that there is something surprising about the appearance of antecedently explored mathematics in physics.

First, making mathematics applicable to a physical phenomenon often involves dramatically simplifying the physics of the situation – ignoring physical phenomena that the physicist knows to be real, and drastically simplifying others – precisely in order that the physicist might then have a chance of successfully applying mathematics to the phenomena that interest him.

Second, even once this is done, it is often only imprecise and non-rigorous versions of antecedently explored mathematical structures – sometimes bearing only impressionistic relationships to any structures or theories to be found in the mathematics journals – that find their way into the language of physics.

These two points show the great extent to which the physicist must struggle, making sacrifices in both the physics (the first point) and the mathematics (the second point), in order to successfully apply mathematics to his discipline. Insofar as this struggle involves so much compromise and sacrifice, there is little reason to think of the applicability of antecedently explored mathematical structures in physics as surprising.

My third point is that when one looks at many of the applications of mathematics in physics, one is struck by the fact that it is often only the more superficial aspects of mathematics – as opposed to the 'deep theorems' of which the mathematician is so proud – that end up finding their way into the physicists' toolbox. Only the shallower, more computationally oriented aspects of mathematics – the sorts of things that a physicist, with little of the mathematician's blood in him, could figure out for himself – are really used in physics. There are some exceptions to this that I will have to discuss later. But insofar as this is generally true, our ability to think of the historical applicability of mathematics as 'unreasonable' or 'surprising' is, I think, greatly diminished. I see little reason, for instance, to be amazed at the applicability of addition or matrix multiplication in physics. Yet the mathematics that one often finds applied in physics is far closer to addition or matrix multiplication than the sorts of things even an advanced graduate student in mathematics would be acquainted with.

If all of this is true, there is little reason to feel wonder at the presence of previously investigated mathematical structure in physics. I shall now consider each of these three points in turn.

7.2.2 Approximations and Idealization.

When trying to understand a phenomenon in physics, one of the first things the physicist must do is to simplify the situation. The physicist will treat a body as a point mass or a perfect sphere, he will treat a spring as a simple harmonic oscillator, and he will treat a hydrogen atom as a light point charge rotating around a heavy, oppositely charged point nucleus.

What dictates the way in which we idealize? One might think that some process like perturbation theory tells us what sorts of idealizations are acceptable, but we have seen in Chapter 2 that perturbation theory, *on its own*, cannot perform this function. More generally, the arguments from Chapter 2 suggest that theoretical considerations (i.e., considerations which *presuppose* some particular theory) cannot account for the ways in which we idealize. Factors of a more 'extrinsic' nature must somehow be at play.

Surely one of the important determinants of the way in which we idealize a physical system is our desire to render the system amenable to some sort of treatment using familiar mathematical techniques. This is by no means a trivial task – if we try, even in classical mechanics, to treat the motion of a physical body in full precision – taking into account all the internal stresses, and so on – what we end up with is an overwhelming and utterly intractable mess of partial differential equations, not all of which may even be properly posed. Progress by known methods becomes impossible – and this is even assuming all the idealizations of classical mechanics. The lesson, then, is that if we are to make any progress in physics, we must idealize with a *deliberate* eye to producing a *tractable* mathematical formalism. This is surely one of the main guiding forces that lies behind idealization quite generally. Most importantly, this sort of idealization often involves dramatically and knowingly simplifying the physics of the situation, in order that we might have a chance of successfully applying known mathematics to the case at hand.[3] [4] It involves, for instance, ignoring internal stresses, and making all sorts of idealized, unphysical assumptions.

This process of idealization is the first way in which the physicist must fight to make mathematics applicable to his discipline. The physicist must carefully and knowingly ignore and idealize many aspects of the physics of the phenomena under investigation before he can even begin a serious discussion with the mathematician. In this way, the applicability of mathematics to physics arises only after a great deal of compromise on the physicist's part.

7.2.3 Non-Rigorous Mathematics.

The physicist, however, must not only perturb the *physics* of a situation in order to render mathematics applicable; often he must also perturb the *mathematics* in order to make it useful. This happens most blatantly when he adopts non-rigorous argumentation in order to make inferences to which he feels independently entitled – but it happens in other ways too.

Several provocative examples of non-rigorous mathematics in physics have been given by Steiner in [90]. Let me also remind the reader of some of the examples discussed in Chapter

[3] An excellent example of this is provided by French, on page 114 of [35]:

> *"The construction of isospin is a classical example of the development of the Wigner programme, ... Here the effectiveness of mathematics surely does not seem quite so unreasonable, as group theory is brought to bear by a series of approximations and idealizations, ... In effect, the physics is manipulated in order to allow it to enter into a relationship with the appropriate mathematics, where what is appropriate depends on the underlying analogy. At the most basic level, what motivates this manipulation and therefore underpins the effectiveness of mathematics in this case are the empirical results concerning intra-nuclear forces and the near equivalence of masses."*

[4] Azzouni makes a related point in [5], page 214:

> *" ... successfully applied mathematics doesn't need a 'perfect' fit between mathematical ontology and empirical phenomena. ... Explaining why a mathematical subject matter can be applied to an empirical domain at all utilizes exactly the same tools to explain why failures in prediction often arise. In particular, the fact that what we need explained is not a perfect fit between mathematics, and what it's applied to, helps make such an explanation possible."*

Azzouni suggests that because we don't need a complete isomorphism between mathematics and reality, we shouldn't be surprised at the usefulness of mathematics. To complete this argument, however, we need to supplement Azzouni's claim with the observation that the 'imperfections of fit' that occur are not mere accidents, but rather the conscious and deliberate choices of physicists, made precisely in order that a mathematical physics be possible.

3. Perhaps the most famous example of non-rigorous mathematics in physics is the use of the delta function in quantum mechanics. In almost all quantum mechanics textbooks, one finds position eigenstates used over and over, even though, of course, such things do not exist. The manipulation of such eigenstates is an indispensable part of many of the arguments given by the pioneers of quantum mechanics – even though the authors of such arguments knew full well that what they were doing was mathematically unacceptable. They used position eigenstates anyway, brazenly ignoring the mathematician's warnings to the contrary.

For another example, note that Feynman's path integral formulation of quantum mechanics assumes the existence of a path integral measure that, in general, does not exist (see [15], as well as Appendix 1.) Nevertheless, the path integral remains widely used in quantum mechanics, especially in quantum field theory, where it provides the *only* language for the discussion of certain phenomena. Again, mathematical concerns about the existence of the path integral have had little effect on the deep reliance of modern quantum mechanics on this particular mathematical tool.

The fact that some concepts such as the delta function subsequently turn out to be rigorizable detracts nothing from my main point. All that matters is the empirical fact that, in rendering mathematics applicable to physics, physicists often find themselves driven to contort the mathematics into a form that most mathematicians would regard as unacceptable. They do this mainly because they see no alternative. Brilliant mathematicians may see such alternatives decades later, but usually only after great intellectual battles are fought and won. The fact that such struggles are so often necessary, and so often difficult, only *helps* make that case that rigorous mathematics – the pride and joy of the mathematician – is often, on first blush, completely unhelpful in making progress in physics.[5] [6]

[5]The fundamentally different attitudes between the mathematician's and physicist's attitude towards such things as rigor were also discussed by Schwartz on p. 357 of [83]:

> *"The mathematician turns the scientist's theoretical assumptions, i.e., convenient points of analytical emphasis, into axioms, and then takes these axioms literally. This brings with it the danger that he may also persuade the scientist to take these axioms literally. The question, central to the scientific investigation but intensely disturbing in the mathematical context – what happens to all this if the axioms are relaxed? – is thereby put into shadow. In this way, mathematics has often succeeded in proving, for instance, that the fundamental objects of the scientist's calculations do not exist."*

[6]Wilson makes a similar point on p. 145 of [102]:

One should note that the physicist often perturbs mathematics to suit his needs in other ways too. Even the best arguments of the physicists are often hopelessly incomplete – for instance, integral signs are exchanged freely without checking the necessary conditions, Taylor series are assumed to exist without question; the list goes on. In general, the physicist shows little respect for the norms of mathematical argumentation when trying to make mathematics relevant to his discipline. This more general type of disrespect shown by the physicist to the norms of the mathematician is characterized nicely by Schwartz on p. 357 of [83]:

> *"The physicist rightly dreads precise argument, since an argument which is only convincing if precise loses all its force if the assumptions upon which it is based are slightly changed, while an argument which is convincing though imprecise may well be stable under small perturbations of its underlying axioms."*

In addition to the sacrifices of physical knowledge the physicist must make, the application of mathematics to physics therefore also demands a sacrifice of many of the most cherished properties of mathematical theories: their rigor, detail, and precision.

7.2.4 The Superficiality of Mathematics.

Finally, I shall discuss the fact that it is often only fairly superficial aspects of antecedently known mathematics that end up finding uses in physics. This greatly undermines the seriousness with which we can take the historical problem.

Mathematicians are most proud of their deepest theorems. Accomplishments like the Hahn-Banach Theorem or the beautiful theorems associated with Galois theory leave the mathematician feeling most satisfied. Yet theorems such as these – the true pride of the mathematician – almost never find any use in physics. It is more often the less striking and more gritty aspects of mathematics that find their way into physics instead.

The concept of a Hilbert space, for instance, is not mathematically deep. Of course, one can prove deep theorems about Hilbert space. But in order to perform any calculation in quantum mechanics that a physicist is likely to want to perform, one can get by with just

"On the contrary, it seems to me more likely that mathematics is characterized by its own internal methods of investigation, which can be adapted only with difficulty to physical circumstances."

the definition of a Hilbert space, and little else. One doesn't need to know *any* theorems about Hilbert space in order, for instance, to calculate the atomic spectrum of hydrogen or helium – including known relativistic corrections.

This sentiment has been echoed most effectively by Zee. Consider his distinction between 'mathematics' and 'arithmetic', given on pages 309-310 of [105]:

> *"... I want to distinguish between mathematics and, for a lack of a better term, arithmetic. This distinction is frankly and intentionally touched with a measure of snobbery. ... I would say that mathematics is whatever a reasonably brilliant physicist, defined for the purpose here as someone significantly smarter than I am, could not work out in a finite amount of time, by following more or less straightforward logic. ... Everything else is arithmetic."*[7]

Zee goes on to argue that much of physics really uses nothing more than arithmetic. To make his case, he enlists the help of Feynman on page 310:

> *"Echoing a fairly widespread arrogance of the physicist, Feynman once said that had God not created mathematicians, physics would have been delayed by about a week. ... According to Feynman, physicists would have invented what they needed, and the rest, as far as he was concerned, should not have been bothered with in the first place. Feynman's attitude, of course, represents a long tradition in physics. "* [8]

On page 311, Zee presents the Standard Model as perhaps the most striking verification of this claim:

> *"In a way, Feynman was right. We reached the grand unified theory in the 1970s using a minimal amount of mathematics. Some fairly elementary group theory, that's about it. For Pete's sake, the grand unified theory! The theory that unifies three of the fundamental interactions! We have unraveled a big piece of Nature's innermost secrets about how the world is put together without using anything a mathematician would call mathematics."*

Further evidence for my claim is also found in following exchange between Dirac and Weyl (taken from page 13 of [22]):

---

[7] Zee continues:

> *"For instance, I probably could have figured out the properties of the solutions of Legendre's equation. All that stuff about Legendre polynomials is definitely arithmetic. On the other hand, the fact that there are only three cases in which higher dimensional spheres can be mapped non-trivially onto lower dimensional spheres, namely $S3 \to S2$, $S7 \to S4$, and $S15 \to S8$, that I call mathematics. ... In short, I associate mathematics with structural or global understanding, and arithmetic with computation."*

[8] See also page 311 of Zee:

> *"Incidentally, Feynman once told me, while we were watching some show, that fancy schmancy mathematical physics as applied to physics is not worth 'a bottle of piss-water'."*

> *"In 1928 Dirac gave a seminar, at the end of which Weyl protested that Dirac had said he would make no use of group theory but that in fact most of his arguments were applications of group theory, Dirac replied, "I said that I would obtain the results with no* ***previous*** *knowledge of group theory!""*

In his talk, one imagines Dirac plowing ahead, doing his calculations in the most natural way possible for a physicist, not thinking of himself as doing anything mathematically subtle, and hence feeling that there was no way that what he was doing should be called anything as sophisticated sounding as 'group theory'. Insofar as what he was doing *could* be seen as invoking group theory, it was surely group theory at a superficial level – indeed, this is surely what Dirac meant by saying that no *previous* knowledge of group theory was presupposed – everything needed could be introduced and absorbed within the lecture! I think that most of the supposedly 'unreasonable' ways in which mathematics is useful in physics are similar to this. The use of Hilbert space in quantum mechanics, the use of group theory in particle physics, and the use of Riemann space in general relativity all occur in a similarly superficial manner.

Some exceptions, however, ought to be noted. While the Standard Model can be developed and explored with only simple mathematics, much of the research that has gone into unified field theories – in particular, quantum gravity – have involved very sophisticated mathematics. Zee explicitly mentions this on p. 310:

> *"Until the mid-1970s, I would have been inclined to agree with Feynman, but with the advent of superstring theory, and for about a decade before that, truly profound mathematics has started coming into physics, with an intensity that was last seen with the arrival of group theory into quantum physics. . . . "* [9]

Does the use of high-powered mathematics in areas of physics such as string theory give new force to Wigner's sense of wonder? I do not think that it does. I will give two reasons for saying this.

[9] See also p. 312 of Zee:

> *"With the shackles of Feynman diagrams broken, Feynman's view on mathematics also started to fade. A younger generation of particle physicists felt increasingly at ease with modern mathematics. There was a fundamental shift in the outlook towards mathematics, and with the advent of superstring theory around 1983 or so, the trend has accelerated. Today, much of the research in superstring theory is really research into mathematical structures, of a degree undreamed of by Wigner."*

The first reason is that the areas of contemporary physics to which high-powered mathematics have been recently applied remain incomplete, tentative, and empirically unconfirmed. We cannot be sure that the use of high-powered mathematics to develop a theory of quantum gravity, for instance, is anything more than a futile attempt to go down a path that nature does not recommend. For all we know, the correct quantum treatment of gravity – if such a thing exists – is mathematically simpler than anyone presently imagines. Consequently, it is premature to judge that Wigner's sense of wonder has been firmly vindicated by the appearance of highly sophisticated mathematics in theories of quantum gravity.

But of course, I do not want to rest my case on this possibility. It may well turn out that the high-powered mathematics used to explore quantum gravity really *is* necessary. This then brings me to my second point – insofar as areas of physics such as string theory rely on sophisticated mathematics, we tend to find physicists participating in the discovery and development of the relevant mathematics. Physicists such as Witten, for instance, have contributed a great amount to areas of mathematics such as K-theory and M-theory that have been judged to be important for research in string theory. (In fact, the mathematical contributions of Witten have been so great that in 1990 he was awarded the Fields medal for his work.) The extraordinary contributions of physicists to areas of mathematics such as these have actually prompted mathematicians to write journal articles discussing how to deal with the somewhat idiosyncratic style of mathematical argument employed by the physicists in their purely mathematical works. (See, in particular, [3] and [46].) That these articles have appeared in prestigious mathematical journals shows how substantial the physicists' contributions to these areas are.

String theory therefore teaches us that insofar as physicists want to apply deep mathematics to physics, they must invent it themselves, or, at the very least, exert a strong force on the mathematicians who they hope will do the work for them. One can also extract this lesson from the history of the calculus, which was devised by Newton with specific physics problems in mind. The general lesson here is that the mathematics community has never been any good at giving the physics community the gifts they want in an unsolicited manner. If the mathematics needed is really difficult, the physicist must prod and poke the mathematician to give him what he wants, or do it himself. Thus, the unusual appearance of sophisticated mathematics in physics, far from vindicating Wigner's sense of wonder, actually helps to dissolve it, insofar as physicists tend to find themselves involved in its development.

The broad claim of this section is that insofar as the physicist takes himself to have received anything useful from the mathematician, it is either something that was trivial, or solicited. If this is right, the sense of mystery with which we are supposed to view the applicability of antecedently explored mathematics in physics largely disappears.

7.2.5 Summation.

When all of these points are duly considered, the role of antecedently discovered mathematics in physical discovery appears somewhat more messy and less magical than is sometimes portrayed. The physicist must neglect much of what he knows, in order to take some of the more superficial structures of mathematics, and apply them in what is often a haphazard, undisciplined and unrigorous way to physics. If he wants to apply sophisticated mathematics, he must often develop it himself. This seems to make it an exaggeration of the highest order to say, as Wigner does, that *"[t]he miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve".* ([101] p. 9.) If there is any sense of mystery to be found with the applicability of mathematics in physics, I do not think it is to be found in the historical problem.

7.3 THE AESTHETIC PROBLEM.

7.3.1 Introduction.

I wish now to consider the second version of the problem of the unreasonable effectiveness of mathematics in physics – the *aesthetic* problem.

Over the course of his education, the good mathematician develops a keen sense of what sorts of mathematical structures warrant serious attention, what sorts of arguments are most natural in a given circumstance, what sorts of shapes it is most natural for a proof to take, what counts as a deep mathematical result, what counts as a triviality, and so on.

This sense, which is gradually instilled in the student of mathematics over a great number of years, involves many judgments which are seemingly *aesthetic* in nature. For instance, judgments concerning what sorts of theorems would be nice if true, and what sorts of proofs would be nice if they worked, form an indispensable part of the mathematician's intellectual sensibility, and are generally taken to reflect the mathematician's aesthetic sense, rather than anything else.

The mathematician's aesthetic sense is an esoteric and rarefied thing. There is little reason to expect that it will have any use outside the context of pure mathematics. We would not, for instance, expect this rarefied sensibility to be especially useful in areas such as political science, psychology or geology. And in many areas where we *might* fantasize about it being useful – such as in predicting the future of an economy, the stock market, or in winning games of chess – the mathematician's aesthetic sense turns out to be of far more limited value than we might have hoped.

Yet when we come to a science like physics, things are remarkably different. By exercising his distinctly mathematical sensibilities, the physicist is often led to make conjectures which turn out to be true. The mathematician's aesthetic sense therefore turns out to be an extraordinarily efficient guide in the formulation of conjectures in physics.

Steiner has argued in [91, 92] that this is a bona fide mystery. He characterizes the problem on p. 47 of [92]:

> *"...[H]ow does the mathematician – closer to the artist than to the explorer – by turning away from nature, arrive at its most appropriate descriptions?"*

Steiner is puzzled by the fact that the mathematician's sense of elegance and convenience have been useful guides for physicists. See p. 7 of [92]:

> *"...relying on mathematics in guessing the laws of nature is relying on human standards of beauty and convenience. So this is an anthropocentric policy; nevertheless, physicists pursued it with great success."* [10]

With all this in mind, Steiner declares on p. 455 of [91]:

[10] Indeed, the entire goal of Steiner's book [92] is to convince the reader that there is a bona fide mystery here. See p. 55:

> *"My topic is anthropocentrism, and my goal in this book is to show in what ways scientists have quite recently and quite successfully adopted an anthropocentric point of view in applying mathematics."*

> *"The historical success of mathematics in physical discovery, on the other hand, does not seem to me to be a coincidence – and therefore demands some sort of account."*

Steiner is not the only person to have considered this particular form of the problem of the unreasonable effectiveness of mathematics. On p. 178 of [71], for instance, Polkinghorne writes:

> *"After all, mathematics is conducted in the human mind. It is the exploration of a rational world within. Yet I am claiming that some of the most elegant patterns encountered in that interior voyage of discovery are found to be realized in the exterior structure of the physical world around us."*

Obviously, things might have been otherwise. We can imagine a world in which the mathematician's aesthetic sense is of no use to the physicist, and the fundamental laws of nature are ugly, mathematically unnatural, and do not exhibit any of the structure found in ordinary pure mathematics. One might even think that this is what we ought to expect – the mathematicians's aesthetic sense is such an idiosyncratic thing that there seems little reason to think that it should be useful to the physicist. Yet, contrary to these expectations, it turns out that the physical world *does* seem to shape itself with the instincts of the mathematician in mind. This is surely a surprise.

The fact that the mathematician's aesthetic standards have proven so useful to the physicist in formulating conjectures, I shall call the *aesthetic problem of the unreasonable effectiveness of mathematics in physics*, or just the *aesthetic problem.*

There are several concerns one might have about the aesthetic problem. In particular, one might wonder whether the way the aesthetic problem has been formulated involves a misrepresentation of the nature of mathematics. Is it really fair to say, for instance, that the mathematician's sense of naturalness or convenience is an anthropomorphic thing? How esoteric *really is* the mathematician's aesthetic sense? I will not pursue this line of thought here, however. To properly address it would require a full-fledged philosophy of mathematics that I cannot construct at present. There is, however, a different way of arguing that the aesthetic problem is a pseudo-problem, that involves thinking more about the relationship between physics and mathematics. It is this class of objections that I shall focus on.

My main objection to the aesthetic problem – much like my objection to the historical problem – is that it is a misrepresentation to claim that the mathematician's aesthetic sense has been a tremendously useful tool in physical discovery. The aesthetic problem, much like the historical problem, rests on a reading of the history of physical discovery that is at once exaggerated and inaccurate.

To see this, first note that some of the criticisms of the historical problem given earlier apply equally well to the aesthetic problem. The aesthetic sense of the mathematician is *not* especially effective, on its own, for discovery in physics, unless coupled with a brazen willingness to ignore some of the physics and many of the norms of mathematics. Successful application of mathematics in physics would therefore seem to rest as much on knowing how to violate one's mathematical sensibilities as it depends on the use of such sensibilities. Furthermore, insofar it is often only mathematics at a fairly superficial level that finds its way usefully into physics, it is only the crassest form of the mathematician's sensibility – rather than anything truly arcane – that the physicist finds useful in forming conjectures. Given all this, it seems a stretch to think of the rarefied aesthetic sense of the mathematician as playing an 'unreasonably effective' role in the process of physical discovery.

In addition to these objections, there are two further objections to the legitimacy of the aesthetic problem that I wish to raise. I discuss these objections not so much because the objections already given are not decisive, but rather because these additional objections illuminate further the awkwardness of the relationship between mathematics and physics.

First, familiar mathematical patterns in new physical theories are often only identified sometime *after* the relevant physical discoveries themselves. In many cases, physicists, in their own clumsy way, unknowingly reinvent an awkward version of the mathematics on their own, leaving it to subsequent generations of thinkers to do their best to interpret their results more elegantly in terms of antecedently known mathematical structures and arguments. The 'surprising' applicability of the mathematician's rarefied sensibility in physics often involves careful and deliberate revisionist storytelling. The mathematician's sensibilities generally play little role for the pioneering physicist – even though they obviously serve as an excellent guide for those whose goal it is to find shadows of the structures and arguments of traditional mathematics in the mass of formulae and techniques that the physicist provides them with. But insofar as it is primarily for this later group that the mathematician's rarefied sensibilities are important, and not the former, the aesthetic problem loses a great deal of its force.

My second objection is that the mathematician's sensibilities are, in a great many cases, just as capable of leading a physicist away from discovery as they are of leading him towards it. I think there is actually little correlation between the pursuit of mathematical elegance or convenience and genuine physical discovery. If this is true, again, the aesthetic problem loses much of its legitimacy.

I shall now say a little more about both of these additional objections.

7.3.2 Historical Revisionism.

Often, it is only *after* physicists chart the territory of an area of physics that it is recognized that their results can be phrased more elegantly with the help of antecedently known mathematical techniques and argument forms. Physicists often show very little sensitivity to the aesthetic sensibility of mathematicians in their initial attempts at physical discovery. Instead, only with careful revisionist storytelling is it possible to motivate their discoveries in a way that appeals to the taste of the mathematician.

The example of Weyl's criticism of Dirac given earlier (taken from Coleman [22], p. 13) concerning the role of group theory in physics is an example of this. In Dirac's lecture, one imagines Dirac following his nose, obtaining his results with no absolutely no thought as to the nature of groups, perhaps even shunning the interference of complicated mathematical machinery where his own sophisticated common sense – as opposed to any arcane mathematical sensibility – seemed to be doing the job perfectly well. Of course, Weyl was surely correct to notice that much of what Dirac had done could be re-interpreted in terms of group theory. But this does not make Dirac's original arguments instances of group theory any more than the interpretation of a two digit addition as a proof in Peano Arithmetic makes the former an instance of the latter. Some may find it curious that the stories of discovery in physics can be retold in a way that titillates the mathematician; but this does not change the fact that the actual context of discovery often involves a process far less attuned to the mathematician's sensibilities than Steiner would have us imagine.

The fact that physicists often proceed in ignorance of the mathematician's refined aesthetic sense and vast store of knowledge was well-known even to Wigner. See pp. 4-5 of

[101]:

> *"It is true also that the concepts which were chosen were not selected arbitrarily from a listing of mathematical terms but were developed, in many if not most cases, independently by the physicist and recognized then as having been conceived before by the mathematician."*

If this is correct, then we should not be amazed at the apparent usefulness of the mathematician's sensibilities in physics. Instead, we should be amazed at the brilliance of those who are capable of seeing the shadows of antecedently known mathematical structures and arguments in the complicated mass of formulae and techniques that the physicist provides them with. When a magician pulls a rabbit out of a hat, a simpleton stares in wonder at the hat, while the informed man is filled with admiration at the magician. Likewise, we should not be amazed by the aesthetic problem; if anything, we should wonder how it is possible to tell a story about so many physical discoveries in a way that appears tailor-made to titillate the mathematician. But to this latter problem, the objections already developed are surely enough to dismiss the sense of wonder.

7.3.3 Is Mathematical Beauty a Good Guide in Physics?

Many physicists have felt that the mathematician's aesthetic sense, although capable of being tremendously helpful to the physicist, is just as capable of misleading him. In fact, some great physicists have explicitly shunned the intrusion of well developed mathematical theories into fledgling areas of physics, feeling that the premature injection of a well developed area of mathematics, with all its aesthetic prejudices intact, into an immature area of physics, can sometimes do more harm than good. In the preface to his *'Treatise on Electricity and Magnetism'* [58], Maxwell tells us:

> *"...before I began the study of electricity I resolved to read no mathematics on the subject until I had first read through 'Faraday's Experimental Researches in Electricity'."*

Why did Maxwell want to read Faraday, and not the mathematicians? Faraday was a largely uneducated man, well known for his ignorance of mathematics. Amazingly enough, Maxwell – a first-rate mathematical mind himself – thought that Faraday's mathematical ignorance was actually an asset. In his essay *'On Action at a Distance'* [59], Maxwell praised Faraday's distinctly unmathematical methods:

"... Thus Faraday, with his penetrating intellect, his devotion to science, and his opportunities for experiments, was debarred from following the course of thought which had led to the achievements of the French philosophers and was obliged to explain the phenomena to himself by means of a symbolism which he could understand, instead of adopting what had hitherto been the only tongue of the learned. ... But Faraday, by a series of steps as remarkable for their geometrical definiteness as for their speculative ingenuity, imparted to his conception of these lines of force a cleanness and precision far in advance of that with which the mathematicians could then invest their own formulae."

Maxwell felt that by being unconstrained by the mathematical prejudices of his time, Faraday was able to understand the physics of electromagnetism in a way that the mathematicians had not been able to. In other words, Maxwell felt that the sensibilities of the mathematicians had been a barrier to progress in electrodynamics, and that, in virtue of being underexposed to the influence of mathematics, Faraday had been given a type of insight that was unavailable to others. Thus it would seem that, in this case, Maxwell felt that the mathematician's aesthetic sensibilities were more capable of blinding the physicist than enlightening him.

Another first-rate physicist – this time, the Nobel laureate Freeman Dyson – similarly expressed his contempt for the intrusion of the mathematician's aesthetic sense into physics, suggesting again that the mathematician's aesthetic preferences are often capable of leading the physicist badly astray. See, for instance, p. 105-106 of [28], where Dyson describes the struggle the pioneering physicist often has to make *against* the mathematician's aesthetic preferences in order to make progress:

"Mathematical intuition is more often conservative than revolutionary, more often hampering than liberating. The worst of all the historical setbacks of physical science was the definitive adoption by Aristotle and Ptolemy of an earth-centered astronomy in which all heavenly bodies were supposed to move on spheres and circles. The Aristotelian astronomy benighted science almost completely for 1,800 years ... the primary reason for the popularity of Aristotle's astronomy was a misguided mathematical intuition that held only spheres and circles to be aesthetically satisfactory. ... When Kepler in 1694 finally demolished the epicyclic cosmology by his discovery that planetary orbits are ellipses, he was not helped by any mathematical preconceptions favoring elliptical motions. On the contrary, he had to fight tooth and nail against his own mathematical prejudices, which were still uncompromisingly medieval. Only after years of struggling with various systems of epicycles did he overcome his conservative tastes enough to consider a system of ellipses. Such mathematical conservatism is the rule rather than the exception among the great minds of physics."

This quote shows the way in which the mathematician's sense of what counts as 'natural' can be untrustworthy. One can almost understand a medieval scholar who might have felt that heavenly bodies simply *had* to move in circles, because the asymmetry of an ellipse was too repugnant a thing to be instantiated in nature. There are almost certainly all sorts of similar prejudices alive in physics today; some probably so ingrained as to be invisible. In many cases, the overcoming of such prejudices is the key to future progress. Thus mathematical sensibilities can sometimes hold the physicist back, diverting him from what great progress often requires – daring acts of nerve in which the physicist abandons deep mathematical prejudices, and rushes headlong into the dangerous unknown instead.

There are other related problems with the use of the mathematician's aesthetic sensibilities in physics. For one, there is no guarantee that the laws of physics be mathematically beautiful, and hence no guarantee that the mathematician's aesthetic sensibility will be of any real use to the physicist in the context of discovery. Pure mathematicians, for instance, are not attracted to Maxwell's equations or Schrödinger's equation because, as equations, they simply are not beautiful enough. The seasoned pure mathematician shudders when he sees these equations, and sees no reason to give them special attention. When the physicist says that Maxwell's equations are beautiful, it is because he has studied the extraordinary way in which such a compact set of equations can explain so many phenomena. But this sense of awe is an entirely different thing from anything belonging to the mathematician's aesthetic sensibility, as the neglect of these equations in the pure mathematics literature shows. The mathematician could not have been steered to these equations by following his aesthetic sense.

Another good example of mathematical ugliness in physics is quantum chromodynamics, whose Lagrangian is a *mathematically* hideous expression into which 19 adjustable parameters must be inserted manually. [11] If our best theories are mathematically ugly in this way,

[11] Polkinghorne discusses the case of QCD on page 175 of [71]:

> *"Quantum chromodynamics is such a theory ... there are 19 adjustable parameters which have to be 'put in by hand' before an attempt can be made to correlate the predictions of the model with experiment. The standard model is phenomenologically successful, but no one feels quite content with it as a candidate for an ultimate theory of the structure of matter. There is a two-fold reason for this discontent. One point is that the theory does not bring within its embrace the fourth basic force observed in nature, namely gravity. ... The other point of discontent is the mathematical ugliness of the standard model – all those groups and all those parameters. ... No scheme has yet been found having about it the majestic mathematical beauty which would encourage the thought that "surely this must be right"."*

as they seem to be, it is difficult to see how much faith can be put in the mathematician's aesthetic sensibility to guide the physicist in the right direction.

7.3.4 Summation.

The aesthetic problem of the unreasonable effectiveness of mathematics presupposes that the mathematician's aesthetic sense is a valuable tool in physical discovery. But history does not support this supposition.

In some cases, we can tell the story of a particular physical discovery in a way that makes it seem mathematically natural or inevitable. But such stories often hide the historical reality in which physicists had to battle the mathematical prejudices of their time in order to make progress. If things such as displacement currents and Lorentz contractions really were as mathematically natural as some have suggested, we must be baffled as to why they required the greatest and most mathematically brilliant physicists of their time, at the height of their powers, to be discovered. In fact, the role of the mathematician's sensibilities in physical discoveries is far less helpful than such arguments suggest. Once this is realized, the mystery surrounding the aesthetic problem largely evaporates.

7.4 OTHER PROBLEMS.

I wish to comment on some other putative mysteries about the unreasonable effectiveness of mathematics in physics, before moving on to the descriptive problem in the next chapter.

Many physicists have used the metaphor that, when they stumble on the correct mathematical formulation of a physical law, they get something out of the mathematics that they did not put in. In this way, mathematics plays an unreasonably effective role in the articulation of the laws of physics. See Feynman [34], p. 171:

> *"It is always easy when you have made a guess, and done two or three little calculations to make sure that it is not obviously wrong, to know that it is right. When you get it right, it is obvious that it is right – at least if you have any experience – because usually what happens is that more comes out than goes in."*

The following famous quote from Hertz, taken from [8], makes a similar point:

> *"One cannot escape the feeling that these mathematical formulae have an independent existence and intelligence of their own, that they are wiser than we are, wiser even than their discoverers, that we get more out of them than was originally put into them".* [12]

Let us call the problem of understanding how mathematics enables the physicist to 'get more out than he put in' – part of which must involve clarifying what the metaphor of 'getting more out than was put in' means in the first place – the *more-out-than-in* problem.

On various ways of understanding the meaning of the problem, one can use points similar to those developed in the previous sections to argue that the more-out-than-in problem is a pseudo-problem. But in this section, I want to make a different criticism. Specifically, I want to argue that the most natural way of understanding the more-out-than-in problem really has nothing to do with either mathematics or physics.

To see why, imagine that, based upon noticing that some finite set of x's have a property Φ, one conjectures that all x's have a property Φ, and that this conjecture turns out to be true. Furthermore, imagine that the fact that *all* x's have property Φ turns out to explain *other* facts which were previously unknown, or regarded as unexplainable. In such a case, there is an obvious way in which one 'gets out' more than one 'puts in'. By looking at a few x's, one has found oneself led into a chain of reasoning that has yielded much more than one might have hoped for, given one's original reasons – however capricious they might have been – for examining a handful of x's. In precisely this way, the method of conjectures and refutations can, in its finest moments, lead to all sorts of 'miraculous' discoveries. This sense of miraculousness, however, has nothing to do with the mathematical language in which the conjectures in question are formed, nor anything to do with the fact that it is laws of *physics* that are in question. This phenomenon is generic to any sort of scientific investigation. Consequently, the more-out-than-in problem does not represent a real problem about the applicability of mathematics to physics.

In saying this, I do not mean to claim that there is nothing mysterious about man's capac-

[12] Wigner also used this way of talking; in particular, see page 6 of [101]:

> *"Nevertheless, the calculation of the lowest energy level of helium, as carried out a few months ago by Kinoshita at Cornell and by Bazley at the Bureau of Standards, agrees with the experimental data within the accuracy of the observations, which is one part in ten million. Surely in this case we "got something out" of the equations that we did not put in."*

ity to formulate true conjectures about the world. To the contrary, I think that philosophers such as Peirce have done a good job of arguing that there is a genuine puzzle surrounding the fact that the laws of nature are discoverable by us at all.[13] So, although the problem is legitimate, insofar as it is generic to the sciences, it is *not* an example of the unreasonable effectiveness of mathematics in physics.

One could try to make the more-out-than-in puzzle more specific by asking why it is that *mathematics* is so helpful for formulating good conjectures in *physics*. But if one is talking about the use of known mathematical *structures* in formulating conjectures in physics, one simply has a variant of the historical problem; if one is talking about the use of mathematical *values*, one has a variant of the aesthetic problem, and if one is talking about the use of mathematical *language*, one has a variant of the descriptive problem. In each case, the discussion pertinent to that problem applies.

The pattern here applies quite generally. There are a host of puzzles purporting to address the applicability of mathematics in physics. Why should the laws of physics be discoverable by us at all? Why should the laws of physics be thinkable by us at all? Why should there be laws at all? Why are the laws of physics so simple when expressed mathematically? In each case, the problems are either generic to the sciences, or are just variants of the historical, aesthetic or descriptive problems, to which the relevant discussion applies. I conjecture that the only problems about the applicability of mathematics in physics that really have to do with the use of *mathematics* in *physics* are the historical, aesthetic and descriptive problems, and perhaps minor variants thereof. I have already discussed the first two, and dismissed them as pseudo-problems; once all three of these problems have been discussed, I can claim to have exhausted the main senses of Wigner's problem.

---

[13] See, for instance, Peirce's [70]:

> *"7.679 It is in this way that science is built up; and science would be impossible if man did not possess a tendency to conjecture rightly.*
> *7.680 It is idle to say that the doctrine of chances would account for man's ultimately guessing right. For if there were only a limited number n of hypotheses that man could form, so that $1/n$ would be the chance of the first hypothesis being right, still it would be a remarkable fact that man could only form n hypotheses, including in the number the hypothesis that future experimentation would confirm. Why should man's n hypotheses include the right one?"*

To complete my discussion I turn to a discussion of the descriptive problem in the next chapter.

8.0 UNREASONABLE EFFECTIVENESS: DESCRIPTIVE.

8.1 THE DESCRIPTIVE PROBLEM.

In this chapter, I wish to discuss the third version of the problem of the unreasonable effectiveness of mathematics in physics – the *descriptive* problem. Unlike the historical and aesthetic problems, I shall argue that we *do* have a right to be surprised at this particular type of efficacy of mathematics in physics. My argument for this will largely revolve around a new argument for the Representation Failure Thesis.

One often hears it said that 'the language of nature is mathematics'. I have discussed this claim already in Chapter 6, citing Pythagoras, Kepler, Gauss, Galileo and Einstein. There, I focused on one thing that might be meant by the claim that the langauge of nature is mathematics – in particular, I took it to mean that trans-temporal nature is isomorphic to a mathematical structure. One of the main lessons of Chapter 6 is that, if this is what we mean when we say that the language of nature is mathematics, then the claim that the language of nature is mathematics is a non-trivial one, insofar as it is possible (and even likely) for mathematics to *fail* to be the language of nature in this sense. To say that the language of nature is mathematics is therefore to make a substantive metaphysical claim about reality.

In fact, the non-triviality of the assertion that the language of nature is mathematics was also appreciated by Feynman, although I shall argue that he meant something quite different from what I meant in Chapter 6. Consider, for instance, what Feynman says on pages 170-171 of [34]:

> *"[when] we have a mathematical theory by which we can compute consequences, what can we do? It really is an amazing thing. In order to figure out what an atom is going to do*

in a given situation, we make up rules with marks on paper, carry them into a machine which has switches that open and close in some complicated way, and the result will tell us what the atom is going to do! If the way that these switches open and close were some kind of model of the atom, if we thought that the atom had switches in it, then I would say that I understood more or less what is going on. I find it quite amazing that it is possible to predict what will happen by mathematics, which is simply following rules which really have nothing to do with what is going on in the original thing. The closing and opening of switches in a computer is quite different from what is happening in nature."

What is Feynman amazed by here? Let us consider two different possible activities. Activity 1: We prepare an atom in a state i, place it in some sort of apparatus, wait a little while, and then measure its final state f. We write down i and f on a sheet of paper, ending up with a page full of data. Activity 2: We pick some real number(s) i, then solve a mathematical problem involving i as a parameter, obtaining some result f. We write down i and f on a sheet of paper, ending up with a page full of data. It turns out that both Activities 1 and 2 lead to the same result – that is, it turns out that the pair i, f is a possible entry on the sheet corresponding to Activity 1 if and only if it is also a possible entry on the sheet corresponding to Activity 2. That two utterly distinct procedures should produce isomorphic results in this way is what (rightly or wrongly) appears like magic to Feynman.[1]

Let us say that 'mathematics is the language of nature' if for any measurable physical system, the set of pairs i, f that result from preparing the system in state i (at some fixed time t) and measuring the subsequent state f (at some fixed time $t' > t$) is the same as the set of pairs i, f that one gets by taking the solution f to some (fixed) mathematical problem involving the parameter i. (Strictly speaking, this assertion is relative to a pair t, t', so to be more precise, we must say that the language of nature is mathematics if the preceding claim is true for all $t < t'$.) Feynman is suggesting that there is something surprising about the fact that the language of nature is mathematics, in this sense. (Note, however, that we are not obviously talking about the language of nature being mathematics in the same sense as

[1]Feynman is not the only person to have found this so. See page 189 of Rosen [78]:

"On the face of it, there appears no reason to expect that purely syntactic operations (i.e., inferences on propositions about events) should in fact correspond to causal entailments between events in the external world. Most if the time, in fact, they do not. Wigner's miracle is that sometimes they do; if we choose our language carefully, and express external events in it in just the right way, the requisite homology appears between implication in the language, and causality in the world described by the language."

in Chapter 6.)

I shall associate this sense of surprise with what I shall call the *descriptive problem of the unreasonable effectiveness of mathematics* – or just the *descriptive problem.* In brief, to say that the descriptive problem is a bona fide problem is to express surprise at the fact that the language of nature is mathematics in this sense – that is to say, it is to express surprise at the fact that by associating physical objects with mathematical structures, and performing purely formal or mathematical manipulations on those mathematical structures, it is possible to know everything that can be known about the future state of a given object. (We shall restate this more precisely later on.) Whether or not this problem rests on a faulty philosophical conception of mathematics and/or physics is what we shall have to examine. Suffice for now to say that this is a third, conceptually distinct, sense in which we might speak of the unreasonable effectiveness of mathematics in physics.

It will be helpful to distinguish a couple of versions of the descriptive problem. Our current best physical theories are mathematical theories. Yet surely *this* fact is not amazing. To see why, it suffices to reiterate some of the observations of the previous chapter. Where we find mathematics successfully applied, we often find a struggle in which much of physical reality, as well as many of the central norms of mathematics, have been sacrificed. When we realize all that is lost in making mathematics applicable to reality, I think any sense of surprise is greatly diminished. Indeed, we might wonder whether the fact that our physics is mathematical is nothing more than a *selection effect* – perhaps modern physics is just the study of physical phenomena *insofar as their dynamics can be mathematically represented.* If this is right, the fact that modern physics speaks the language of mathematics is no source for amazement at all. The *contemporary descriptive* problem – that is, the problem of why the language of our *best current* physics is mathematics (in the sense described above) – is not a genuine problem.

However, it is also an article of faith amongst many that the *ultimate* structure of the universe is mathematical – where by this, I mean that we can discover the future states of physical systems by solving math problems, in precisely the way described above. In this case, what is imagined is not just that mathematics is the language best suited for *us* in our current discussion of physical nature, but rather that this language is the language that *nature* speaks. The fact that it is *in principle* possible to predict the time evolution of physical systems by solving math problems is what (rightly or wrongly) amazes Feynman. Let us call the problem associated with this source of amazement the *ideal descriptive* problem.

The objections given to the contemporary descriptive problem do not apply to the ideal descriptive problem. It is often imagined that, in the final theory of physics, we can describe the evolution of any physical system in a way that is exact, mathematically rigorous, and in which no physical phenomena are ignored, however insignificant. In such a case, nothing has to be sacrificed in making mathematics applicable to physics. In this sense, physics ultimately reduces to mathematics. The thought behind the ideal descriptive problem is that there is something surprising about this.

From here onwards, we consider only the ideal descriptive problem. I think that the ideal descriptive problem *is* a legitimate philosophical problem, to which none of the objections previously discussed apply. But more still needs to be said about what makes it a bona fide source of amazement – so much so that it managed to provoke commentary from Feynman, a physicist otherwise notorious for his impatience with philosophical questions. Might not Feynman's sense of wonder rest on a philosophical mistake? We shall have to consider this possibility carefully.

In section 2 of this chapter, I wish to consider some preliminary objections that might be given to the ideal descriptive problem. In section 3, I will then give a positive argument in support of the ideal descriptive problem – that is to say, an argument that the ideal descriptive problem *does* indeed present us with a type of efficacy of mathematics in physics that could not have been anticipated. This will involve giving a new argument for the Representation Failure Thesis – specifically, it will involve arguing that it is likely that the language of nature is *not* mathematics, in the sense just described (as opposed to the sense described in Chapter 6). In section 4, I will then present some concluding remarks.

8.2 CRITICISMS OF THE PROBLEM

Feynman is surprised at the fact that problem solving in physics should reduce to problem solving in mathematics. In this section, I would like to assess various arguments that one might imagine being given to rid someone of such a sense of surprise.

Objection 1: Philosophers have often wondered how it is that the abstract concepts of mathematics are applicable to physical reality. While this question is especially challenging to Platonistic philosophies of mathematics, it is a pointed question for many other philosophies of mathematics as well. A common response to this challenge is to say that mathematical concepts are often abstracted from physical reality in the first place. See, for instance, page 194 of Nagel [64]:

> *"It is no mystery, therefore, that pure mathematics can so often be applied. Many branches of pure analysis, it is evident from the preceding historical survey, are developed on the basis of "suggesting sciences", whose fundamental notions, even though highly abstract, are yet drawn from observation of the natural world. It is a reasonable hypothesis that pure mathematics in general is so often applicable, because the symbolic structures it studies are all suggested by the natural structures discovered in the flux of things. Their symbolical statement and development liberates us from the press of local circumstance, and permits us to extrapolate invariant relations found in one segment of the flux to another."*

Let us call any conception of mathematics according to which mathematical concepts are abstracted from physical reality an *abstractionist* conception of mathematics.

If an abstractionist conception of mathematics is correct, then Feynman's sense of wonder is undermined. For if mathematical concepts are suggested to us by an examination of 'the flux of things', then surely we have *every* reason to expect that problem solving in physics should reduce to problem solving in mathematics. If an abstractionist conception of mathematics is right, it would surely be surprising if the language of physics turned out to be anything *other* than mathematics.

Reply: One way of replying to this is to challenge the abstractionist conception of mathematics. Interestingly enough, this is what Wigner does in [101]:

> *" . . . whereas it is unquestionably true that the concepts of elementary mathematics and particularly elementary geometry were formulated to describe entities which are directly suggested by the actual world, the same does not seem to be true of the more advanced concepts, in particular the concepts which play such an important role in physics, . . . "*

Wigner specifically mentions complex numbers, algebras, linear operators and Borel sets as examples of mathematical concepts that do not seem to have been abstracted from physical reality, and that pose the greatest problems for the abstractionist.

A full criticism of the abstractionist conception of mathematics would take us too far afield, into areas of philosophy of mathematics that have little to do with the main points of this chapter. I shall therefore focus on replies to the objection that do not take issue with the abstractionist conception of mathematics. Suffice it to note that an abstractionist conception of mathematics cannot simply be taken for granted.

I shall present two main replies to the abstractionist objection to Feynman's sense of wonder, neither of which challenge the abstractionist conception of mathematics itself. First – it does *not* follow immediately from the fact that mathematical concepts are abstracted from physical reality, that we have any reason to expect that the language of *physics* is mathematics. When someone like Nagel discusses the 'flux of things' from which mathematical concepts can be abstracted, it seems that he means the flux of everyday experience – we notice that apples can be counted and added, we notice that our experience of time seems to have the structure of linear ordering, and from this we abstract certain mathematical concepts. But facts about putting groups of apples in a basket, although pertinent to *physical* reality, are not *physics* by any stretch. Almost nothing about the fundamental laws of physics can be gleaned from the behavior of apples, or from the fact that our psychological perception of time is linearly ordered. So even if the concepts of mathematics are abstracted from physical reality, they must be abstracted from the *right* aspects of physical reality in order for the abstractionist to be able to argue that we have some special reason to expect such concepts to be useful in *physics*.[2] [3]

[2]Of course, it is true (let us assume) that all aspects of reality may be subsumed under the laws of physics. This, however, is *still* not enough for us to reasonably expect that concepts abstracted from our day to day life will be useful in physics. To see why, note that when it comes to most daily events (e.g.: a bus is late, a friend changes his mind about something), all we can do is *conjecture* that somehow such events may be subsumed under the laws of physics. The fact then that a certain concept can be abstracted from our dealings with friends or busses therefore gives us very little reason, on its own, for reasonably hoping that such a concept will be useful in fundamental physics.

[3]Much philosophy of mathematics is concerned with explaining how it is that mathematical concepts and arguments can be applicable in physical reality. The first chapter of Steiner's [92], for instance, is devoted to a discussion of this problem. This area of philosophy of mathematics, however, has nothing to do with questions about the interaction of physics and mathematics. In this area of philosophy, one finds philosophers concerned with understanding how it is that arguments like 'there are 2 apples on the table, there are 5 bananas on the table, 2+5=7, therefore there are at least 7 items on the table' can be construed as valid. But this has nothing to do with physics. The fact that such arguments may or may not be able to be construed as valid under a given philosophy of mathematics tells us nothing about whether the basic laws of physics are mathematical or not. Even when such philosophers assure us that their pet philosophy of mathematics guarantees that electrons can be individuated and counted, such assurances give us no reason to think that

In order for the abstractionist to argue that the mathematical describability of fundamental physics is unsurprising given the genesis of mathematical concepts, he must be able to argue that the mathematical concepts used in fundamental physics were originally abstracted not just from physical reality, but from considerations quite close to fundamental physics. I think, however, that such a claim does not gel well with the history of mathematics.

In order not to get too involved in discussions about the history of mathematics, however, let us present a second reply to the abstractionist's criticism of Feynman, again not involving a challenge to abstractionism itself. Indeed, let us assume that all the concepts of mathematics *are* abstracted from considerations having to do with *physics*, as opposed to considerations merely having to do with *physical nature* (such as apples and tables and chairs). I claim that it *still* does not follow that we have any right to expect that all problem solving in physics should reduce to the solving of purely mathematical problems.

To see why, let us consider an analogy. Imagine a person who enjoys staring at clouds, and comes up with a novel vocabulary for classifying clouds and the way they move. Let us call his set of cloud concepts C. Now, it is certainly possible that many events involving clouds can be described and explained quite adequately with reference to concepts from C – for instance, in certain cases, it may be possible to explain the way a cloud's shape changes over time by noticing that it was a p type cloud, where p is a concept in C, and that p clouds tend to transform into q clouds, where q is also a concept in C. However, although all the concepts in C were abstracted from staring at clouds, and although the concepts of C are rich enough for many purposes, it may well turn out that the concepts of C are *not* sufficient for describing all cloud-events. Indeed, this claim can be generalized: it does not follow from the fact that a set of concepts C is abstracted from some set S of events, that the set of concepts C is expressively rich enough to develop a complete theory about the events of S. We cannot draw conclusions about the *expressive power* of a set of concepts solely from their *genesis*. But surely this is just what the abstractionist in the objection wants to do – he wants to infer from the fact (granted for the sake of argument) that mathematical concepts are abstracted from physics considerations, that we should not be surprised that

the use of mathematical language will be necessary in articulating the fundamental physical mechanisms by which electrons evolve in time. Thus, the literature surrounding such issues is not directly relevant to the question at hand.

the laws of physics can be stated in *purely* mathematical terms. Now, if one has *independent* grounds for thinking of the concepts of mathematics as expressively rich enough to handle all of physics, then that is fine. But the truth of abstractionism – which is a theory about the *genesis* of mathematical concepts, cannot count on its own as such a reason. Abstractionism, therefore, is not (on its own) a threat to Feynman's sense of wonder.

Objection 2: Feynman's sense of wonder revolves around the surprising profitability of treating physical objects as something they are not – namely, mathematical objects. The fact that one can learn something about an object by treating it is something that it is not seems (rightly) surprising to Feynman. But there is another way of thinking about the metaphysics of mathematics according to which nothing is treated as anything it is not. Specifically, we can think of the physicist as discovering the mathematical *properties* that physical objects have, which are then used profitably to form a more complete picture of the object(s) in question – including their future states. The thought that some sort of 'metaphysical confusion' between physical and mathematical objects is necessary whenever mathematics is used in physics is a mistake. Once this is realized, Feynman is deprived of any reason for amazement.

Reply: This objection misidentifies the source of Feynman's wonder. Whether or not one chooses to say that a physical object is being treated as a mathematical object, or that the mathematical properties of a physical object are being exploited, is besides the point. Feynman's sense of wonder comes from the fact that it is *solely* by exploiting the mathematical properties of physical objects that one is able to deduce (in principle) everything that there is to be known about those objects.

Ordinary objects have all sorts of non-mathematical properties; such as color, texture, goodness, and so on. None of these non-mathematical properties occur in the basic laws of physics, however, because they do not need to – it turns out that, in order to understand a physical system at a basic enough level, it suffices to understand only its *mathematical* properties. This is what amazes Feynman. Clearly, things might have been otherwise – in particular, these non-mathematical properties may have turned out to have played a vital role in articulating the laws of physics. And yet, they do not. To do physics, it suffices to exploit the purely mathematical properties of objects. Feynman's sense of wonder therefore boils down to an empirical fact about modern physics, rather than any semantic theory about the interaction of the ontologies of mathematics and physics.

To this, one might object that non-mathematical properties such as color *reduce* to mathematical properties, and that in this sense, even non-mathematical properties turn out to have mathematical underpinnings. Reductionism of this sort therefore seems to deprive us of the distinction between mathematical and non-mathematical concepts that we have used in our reply to the present objection. But this is precisely the phenomenon that amazes Feynman – that we can systematically eliminate the non-mathematical language from our vocabulary, and treat a physics problem as a purely formal mathematical problem. The reduction of the physical to the mathematical, far from eliminating Feynman's sense of wonder, underwrites it.

Regardless of whether Feynman's sense of surprise is ultimately justified, I do not think it can be criticized for resting on an idiosyncratic view of the semantics of mathematical attribution.

Objection 3: Feynman's amazement revolves around the fact that symbolic manipulations suffice to predict the future state of a physical system. However, there is a similar phenomenon when dealing with arithmetic – one can assign cardinalities to distinct sets, go away and perform a purely symbolic calculation, and obtain a prediction as to the result that will be obtained if the aggregate of the two sets is counted. By his own criteria, Feynman should also be amazed at this. But this sense of amazement is surely unwarranted, and so this must make us suspect that his original sense of surprise rests on some sort of philosophical mistake.

Reply: I see no reason to think that Feynman would be amazed at the efficacy of symbolic calculations in revealing *necessary* connections between properties of objects. (One might wonder why the theory of necessary connections is so amenable to symbolic treatment, but that is not Feynman's problem.) What Feynman is puzzled by is the efficacy of symbolic calculations to predict *contingent* connections between properties of objects. He thinks – rightly or wrongly – that it is surprising that symbolic manipulations are so useful in revealing the contingent connections that hold between objects at different times. Questions about the efficacy of arithmetic lie well outside his target.

Objection 4: In the quote we have been examining, Feynman imagines a computer calculating the trajectory of a particle. He notes that the sorts of things going on inside the computer, such as the opening and closing of particular switches, have nothing to do with anything going on in the particle and its environment. This seems to underlie his sense of wonder at the effectiveness of mathematics in physics.

However, whenever an algorithm is performed on a computer (or by hand), we can describe what happens on at least two levels. First, we can focus on what happens at the *physical* level – scribbles are made on a sheet of paper, registers in a computer are altered, and so on. But we can also describe what is happening on a *computational* level – an integral is being calculated, a differential equation is being solved, and so on.[4]

Now, it might be claimed that much of what happens at the physical level of a calculation has nothing to do with the particles whose trajectories are being calculated. For instance, the motions made by my hand in a pencil and paper calculation have little to do with the trajectories of the objects in which I am interested. But one cannot simply conclude from this that the theory of differential equations – which is, let us assume, what I am scribbling about – also has little to do with the trajectories of these particles. Unless Feynman can justify the inference from irrelevance at the physical level to irrelevance at the computational level – which he likely cannot – his sense of wonder rests on a mistake. Surely, after all, there is very little significance in the fact that, at the physical level, the processes involved in the mathematical calculations of physics have little to do with the particles themselves.

Reply: This objection misidentifies Feynman's claim. Although, in the quote, Feynman dwells on the physical level of the calculation, I see no reason to think that Feynman is trying to *argue* for the irrelevance of what happens at the computational level by exploiting what

[4]This distinction is closely related to Dennett's distinction between the physical stance and the design stance, as presented in [23].

he takes to be the irrelevance of what happens at the physical level. I think that Feynman (rightly or wrongly) takes it to be *independently* surprising that, at the *computational* level, symbolic manipulations turn out to be relevant to the trajectories of bodies through space and time. This sense of surprise is *not*, however, the product of an argument that begins with the irrelevance of what happens at the physical level.

Of course, in saying this, I have not explained why Feynman thinks it is independently surprising that, at the computational level, symbolic manipulations turn out to be relevant to the trajectories of bodies through space and time. That will come shortly. But, as long as that promise is kept, the objection in question will have been answered. [5]

Objection 5: There seems no reason why we could not view distinct areas of physics as constituting distinct subject matters. For instance, statistical mechanics and thermodynamics appear to be distinct subject matters, as do the theory of the electromagnetic force and the theory of the weak nuclear force. However, in both cases, a type of unification has been possible – it turns out that one can use the language of statistical mechanics to describe thermodynamics, and one can use the language of the Salam-Weinberg theory to describe, in a unified way, both the electromagnetic force and the weak nuclear force. Now, if Feynman is right to be amazed at the fact that solving problems in physics turns out to be reducible to solving problems in mathematics, then surely he should be equally surprised that, for instance, solving problems in thermodynamics reduces to solving problems in statistical mechanics – in both cases, distinct sorts of activities turn out to be equivalent. If Feynman's sense of wonder is appropriate, we must be amazed every time such a reduction or unification occurs in science. But surely such a sense of amazement is inappropriate – it might be said, after all, that it is the *job* of physics to unify disparate fundamental physical phenomena. A sense of amazement that occurs whenever physics simply does its job is not a sense of amazement that Feynman can expect us to take seriously.

Reply: I do not think that Feynman is committed to being amazed every time some sort of

[5] Admittedly, Feynman's presentation of the problem is a little misleading, insofar as his example focuses on the *physical* level of a mathematical calculation. However, given that our goal is not Feynman exegesis, but rather to elaborate a specific point of view about the relationship between mathematics and physics, it suffices to point out this problem, correct it, and move on.

reduction or unification is achieved in physics. Let us consider the reduction of thermodynamics to statistical mechanics. In this case, it can be said that what we *thought* to be two distinct subject matters – thermodynamics and statistical mechanics – turned out *not* to be two distinct subject matters. Thus, as I have presented Feynman's sense of surprise, I see no reason to think that he must be amazed about either the reduction of thermodynamics to statistical mechanics, or the unification of electrodynamics with the weak nuclear force.[6]

Now, if it were possible to argue that mathematics and physics were not distinct – specifically, if it were possible to argue that there are laws of physics that are logical consequences of the laws of mathematics, or something like that – then Feynman's sense of wonder would not be justified. In such a case, Feynman would have *no* reason to be surprised at the (putative) reduction of physics to mathematics, much as we have no deep reason to be amazed at the reduction of thermodynamics to statistical mechanics (at least, by the arguments presented.) But absent such an argument, Feynman's sense of surprise is left unharmed – there remains a large difference between the person who is mistakenly amazed at the reduction of thermodynamics to statistical mechanics, and Feynman's sense of wonder at the (putative) fact that the language of physics is mathematics.

Objection 6: Given the success with which mathematics has been employed in physics, it is surely reasonable to conjecture that problem solving in physics should always reduce to problem solving in mathematics. The laws of quantum mechanics, relativity, thermodynamics, and so on, are all mathematical laws; in fact, we know of no area of physics requiring laws that *cannot* be mathematically stated. All this surely counts as inductive evidence in support of the mathematical nature of physics. But if we have good inductive grounds on which to claim that the ultimate laws of nature are mathematical, then we cannot say that there is anything *surprising* about the fact that these laws are mathematical, contrary to what Feynman says.

Reply: I do not think that the evidence for the mathematical nature of contemporary physics

[6] This is not to deny that it might be possible to argue, by totally different means, that the examples of scientific reduction and unification given in the objection are 'amazing' in some *other* sense. Such a sense of amazement, however, would be distinct from Feynman's sense.

cited in the objection (e.g., the mathematical nature of quantum mechanics, relativity, and thermodynamics) counts as good inductive evidence for the claim that the *ultimate* laws of physics are mathematical. The considerations of the previous chapter, as well as the considerations presented in my distinction between the contemporary and ideal descriptive problems, help to make this case. The success of mathematics in contemporary physics occurs only after significant sacrifices in both physics and mathematics. The fact that one has had a few successes in forcing an awkward mathematical formalism onto an incomplete description of reality might provide us with hope that we could do better if we were smart enough – but hope and inductive evidence are two quite different things.

I think that all this demonstrates that Feynman's sense of surprise does not rest on any obvious philosophical mistake. In the next section I shall move on to present an argument of a more positive nature in defense of Feynman's sense of surprise.

8.3 DEFENDING THE DESCRIPTIVE PROBLEM.

8.3.1 Analysis.

Feynman is amazed at the fact that solving physics problems might always reduce to solving mathematical problems. Is this anything to be amazed by? This is the question I would like to consider for a while.

What does it mean to solve a mathematical problem as a means of solving a physics problem? Let us consider a special case. Suppose that we are concerned with a deterministic one particle system and how it evolves over time. Let us furthermore suppose that all we need to know about the particle at any moment is captured by some n-tuple of real numbers – for instance, the position and momentum co-ordinates. Using the fact that there exist (definable) bijections $f_n : \mathbb{R}^n \to \mathbb{R}$, we can in fact suppose that all the information about the system is encoded by a *single* real number; i.e., $n = 1$. The way in which the one-particle system evolves over time is then captured by a family of functions $\{\tau_{t,t'} : t < t', t, t' \in \mathbb{R}\}$, where $\tau_{t,t'}(i) = f$ just in case the system in state i at time t will evolve into the state f at time t', where $i, f, t, t' \in \mathbb{R}$. For my argument, it will suffice to consider the physics problem of understanding the way in which our one particle system evolves from $t = 0$ to $t = 1$ – i.e., understanding the process captured by $\tau_{0,1}$.

What might it mean to reduce this particular physics problem to a mathematical problem? For our purposes, a 'mathematical' problem is some formalism (perhaps a differential equation, perhaps something else) that defines a relationship between i and f, where i is the state of the particle at time $t = 0$, and f is the state of the particle at time $t = 1$. So for instance, perhaps f is just the (unique) solution to a given equation involving i as a parameter (e.g; f is the solution obtained at $t = 1$ by solving a differential equation subject to the constraint given by i at $t = 0$.) Importantly, the relation set up between i and f is *definable*; i.e., can be written down as an equation or some other piece of mathematical formalism that can be manipulated by the mathematician.

Now, one might think that *however* our system evolves from $t = 0$ to $t = 1$, it must be able to be captured by *some* mathematical formalism. For instance, let $S_{0,1} = \{< i, f >: i \in \mathbb{R}, f = \tau_{0,1}(i)\}$. Then the system will evolve from i to f just in case $< i, f > \in S_{0,1}$. In this case, the physics problem of deciding whether the system will evolve from i to f is reduced to the purely mathematical problem of deciding whether $< i, f > \in S_{0,1}$ – i.e., a purely mathematical problem about membership in $S_{0,1}$.

Of course, the argument of Chapter 6 shows that things are not so simple – we cannot (in general) presuppose that the set $S_{0,1}$ even exists. But I would like to put such considerations to the side, and pursue an entirely different line of thought here. Let us suppose that the set $S_{0,1}$ is a bona-fide set. Would Feynman be amazed at the fact that it is possible to answer questions about the time evolution of our one-particle system by answering a question about set membership in $S_{0,1}$?

I think that the answer is no. The reason is that all the information about the possible ways in which our system can evolve from $t = 0$ to $t = 1$ has been encoded into $S_{0,1}$ by *brute force* – the 'mathematical problem' of deciding whether $< i, f > \in S_{0,1}$ simply amounts to extracting information from some pre-assembled 'database'. There is nothing amazing about this, because the 'mathematical problem' to which our physical problem has been reduced involves a *parameter* $S_{0,1}$ that has all the information we need directly encoded into it.

Of course, if $S_{0,1}$ were *independently* definable – e.g., if $S_{0,1}$ were the set of roots of some equation, or something like that – then Feynman's sense of wonder would be restored, because our physics problem *would* reduce to a genuinely mathematical problem of finding, for instance, the smallest root of some function involving the parameter i. In such a case, the mathematical problem to which our physical problem has been reduced *could* be stated without reference to *any* parameters (assuming that the definition of $S_{0,1}$ involved no parameters.) *This* is the sort of situation that would amaze Feynman.

In light of this, one might suggest that the physics problem of determining the evolution of our one particle system from $t = 0$ to $t = 1$ reduces to a genuine mathematical problem only if $\tau_{0,1}$ is definable without parameters. However, only countably many functions $\tau : \mathbb{R} \to \mathbb{R}$ are definable without parameters. By contrast, the set of all possible functions $\tau : \mathbb{R} \to \mathbb{R}$ has cardinality $(2^{\aleph_0})^{(2^{\aleph_0})} = 2^{2^{\aleph_0}}$. Thus, any measure over the set of possible time-evolutions of our one-particle system that assigns measure 0 to each point will also assign measure 0 to the entire set of definable time-evolutions. In this sense, it is an anomaly when $\tau_{0,1}$ is definable – i.e., when the physics problem in question is reducible to a genuine mathematical problem. Thus, Feynman is right to be surprised when this turns out to be the case.

While I would like to endorse the spirit of this argument, I would also like to make a technical modification. The solution of mathematical problems associated with physics problems often involve the use of real parameters. For instance, the classical theory of gravitation requires a parameter G, classical electrodynamics requires the parameters μ_0 and ϵ_0, quantum mechanics requires $\hbar$, and so on. We have no reason to think that these parameters are mathematically definable quantities. Therefore, we might allow finitely many real parameters in any definition of $\tau_{0,1}$. However, only $2^{\aleph_0}$ many functions $\tau : \mathbb{R} \to \mathbb{R}$ are definable from finitely many real parameters, out of $2^{2^{\aleph_0}}$ many functions $\tau : \mathbb{R} \to \mathbb{R}$. But by Cantor's Theorem, $2^{\aleph_0} < 2^{2^{\aleph_0}}$. In this sense, the set of time-evolutions $\tau_{0,1}$ that can be obtained as the solution of a mathematical problem involving only finitely many real parameters are vastly outnumbered by the set of time-evolutions $\tau_{0,1}$ that are not. So even in this weaker sense, Feynman's sense of surprise is vindicated.

In fact, we can even allow κ many real parameters in our definition of $\tau_{0,1}$, where κ is any cardinal such that $2^{\kappa} < 2^{2^{\aleph_0}}$ – in particular, this is always satisfied by $\kappa = \aleph_0$. (By including κ real parameters within a first-order language I mean that one has a single parameter

S which is a well ordered set of κ reals.) This is because the number of sets of reals of cardinality κ is $\leq (2^{\aleph_0})^\kappa = 2^\kappa$, and so the number of functions $\tau : \mathbb{R} \to \mathbb{R}$ definable from κ many real parameters is just 2^κ, which is less than $2^{2^{\aleph_0}}$ by assumption.

It follows from this that if $\kappa < 2^{\aleph_0}$ entails $2^\kappa < 2^{2^{\aleph_0}}$,[7] then the set of functions $\tau_{0,1}$ that are definable with the use of a set of reals of cardinality $< 2^{\aleph_0}$ has cardinality $< 2^{2^{\aleph_0}}$. From this we can conclude that, in such a case, most sets of reals can only be defined with the use of a parameter consisting of continuum (i.e., $2^{\aleph_0}$) many reals. But $\tau_{0,1}$ contains $2^{\aleph_0}$ bits of information. So under the assumption just given, most sets of reals X cannot be defined with less information than is contained in X itself.

One should also note that our argument establishes a result a little more general than that suggested in the Feynman quote given earlier. There, Feynman expressed amazement at the fact that some *specific mathematical procedure* took us from i to f, where he imagined such a procedure being performed on a device such as a computer. This would seem to suggest that there be some sort of *recursive* connection between i and f. In my argument, I have only assumed that the relation between i and f is *definable* (possibly involving certain types of parameters), which is a much weaker assumption.

What conclusions can we draw from this? In contemporary physics, the initial and final states i and f of a system can generally be connected by some definable mathematical relation – Maxwell's equations, Schrödinger's equation, and so forth. We cannot assume that this will always be the case. In fact, *most* ways in which a system can evolve defy mathematical description, in the sense that there is *no* definable relation between i and f making reference to only a countable number of real parameters. Thus, we cannot count on the existence of simple (or even complex) mathematical equations that describe the exact and fundamental workings of nature. While we can still sometimes use mathematics to describe such a system (by saying, for instance, that the system will evolve from i to f just in case $<i, f> \in S_{0,1}$), such equations involve at best more information that could ever be used by a human, or at worst as much information as would be contained in a database that simply listed all possible time evolutions by brute force.

[7]This obtains, for instance, if $2^{\aleph_0} = \aleph_1$ – i.e., if the continuum hypothesis is true.

8.3.2 Conclusions.

In the previous subsection, I have argued that Feynman's sense of surprise at the particular way in which mathematics can be used in physics (assuming, that is, that this usefulness carries over to the final theory of physics) is genuinely surprising. In this sense – and, so far as I can tell, only in this sense – Wigner's sense of wonder at the usefulness of mathematics in physics is justified. Thus, in very particular senses, Wigner and Feynman have been defended.

However, we can also use the considerations of the previous subsection to provide us with a new argument for the Representation Failure Thesis. Recall:

The Representation Failure Thesis: We cannot take for granted that the world is representable in mathematical terms – that is to say, we cannot take for granted that the state of the world, together with all its fundamental laws, can be captured perfectly by some set of mathematical assertions or equations. In fact, it is more likely that the world is *not* representable in mathematical terms (in this sense) than that it is.

If the fundamental mechanisms of the world can be captured by a set of mathematical assertions or equations, then there must exist some definable relation between the initial and final states i and f of any system over any period of time. (We allow for the existence of countably many real parameters in any such definable relation.) However, I have just argued that this cannot be taken for granted – in fact, it is only in the minority of cases that the language of nature is mathematics in this sense. The Representation Failure Thesis follows from this.

It is instructive to put this result together with the result of Chapter 6. The modern physicist looks at his toolbox of equations, and conjectures that some set of equations, perhaps similar in form to the ones he has, must be able to be postulated exactly of nature at its fundamental level. But this is not necessarily so. Nature could fail to be isomorphic to any sort of mathematical structure, as shown in Chapter 6. Only a minority of possible worlds are mathematically describable in this way. But even if the language of nature *is* mathematics in this sense, it still might turn out that there is no *definable* relation between the initial and final states of physical systems in the world, and thus there is no *exact* set of equations similar in form to the ones he has written down. In fact, only in a minority of cases do such definable relations exist. The sort of world that our speculative physicist imagines – one in which we may use mathematical equations to capture the fundamental mechanisms of nature – forms a minority within a minority.

8.4 FINAL REMARKS.

I have argued that, a-priori, we have no reason to expect that mathematics be the language of nature (in the sense described in this chapter.) I have even argued that we have some reason to expect mathematics *not* to be the language of nature, insofar as the set of mathematically indefinable time evolutions of a single particle has greater cardinality than the set of mathematically definable time evolutions.

One might accept this, yet still wonder whether Wigner and Feynman's sense of surprise is overblown. Wigner talks of the usefulness of mathematics in physics as not just 'surprising', but 'unreasonable', and even 'miraculuous'. In the quote given earlier, Feynman describes the use of mathematics as 'amazing'. Do the arguments I have given justify such descriptions, or do Feynman and Wigner overstate their cases with such rhetoric?

I do not see any reason for thinking of the effectiveness of mathematics in physics as 'unreasonable', or 'miraculuous'. I suspect that Wigner's use of such terms demonstrates his attachment to something like the historical problem of the unreasonable effectiveness of mathematics in physics, in which mathematicians display a certain inexplicable 'prescience' in the selection of their objects of study. Prescience is often regarded miraculous, amazing, and perhaps even unreasonable. But I have argued that the historical problem rests on a misreading of the history of physics and mathematics, so this way of trying to defend Wigner's rhetoric will not work.

Perhaps the most modest statement Wigner makes is that there is something 'surprising' about the effectiveness of mathematics in physics. I would like to defend this claim, but nothing stronger. That is to say, I would like to argue that the usefulness of mathematics as exhibited in the ideal descriptive problem is 'surprising' – but I will not defend the use of any stronger adjective.

I have argued that the set of possible worlds in which the language of nature is mathematics (in either the sense of Chapter 6, or this chapter) is dwarfed in magnitude by the set of possible worlds in which the language of nature is not mathematics. But this, on its own, does not make it 'surprising' that the language of nature turns out to be mathematics.

To see why, consider the fact that the speed of light is 3.00×10^8 m/s. Presumably, this fact obtains only in an infinitesimally small fraction of possible worlds. Yet any surprise at the fact that the speed of light turns out to be 3.00×10^8 m/s is surely misguided – the speed of light has to have *some* value, and so there is surely something confused about being surprised at the fact that the speed of light turns out to have the value that it has. For another example, imagine writing down a list of all the outcomes of the roulette wheel at a given night in a given casino. The probability that *those* outcomes would have occurred is close to 0. And yet there would be something mistaken about expressing surprise at the contents of the list.

But consider a different pair of examples – imagine that the digits of the speed of light or the roulette list consisted of the digits of pi, in order. This *would* come as a surprise. So some low probability events are surprising, while other equally low probability events remain unsurprising. How can we explain the difference?

First, let me address a concern. One might insist that there is nothing *really* surprising at the list of outcomes at the roulette wheel consisting of the digits of pi – after all, this occurrence is just as probable as any other, and there is no decisive reason to think that it cannot occur. If it occurs, it is a brute fact that simply must be accepted. I do not find this sort of viewpoint very satisfying, however. It is certainly possible to put oneself in a Stoic frame of mind where *nothing* is surprising – after all, if something is the case, there is obviously no decisive reason for it *not* to be the case, and if there is no decisive reason for something *not* to be the case, then we ought not be surprised that it *is* the case. But such a frame of mind is certainly not the everyday frame of mind of any ordinary human being. Nor is it the frame of mind of any scientist – the scientist is constantly pushed and pulled by his conception of what would be reasonable, surprising, amazing, or unbelievable. We must accept, then, that in ordinary life, as well as in science, we can find a fact surprising absent decisive reason for thinking that it not be the case. It is from such a point of view that we must try and assess Wigner's assertion that the effectiveness of mathematics in physics is surprising. Certainly, from a different point of view, Wigner's claim is false – as is any claim of the form 'X is surprising' – but this is not a point of view worth much attention if our goal is to understand whether it is reasonable for *us* to be surprised by some particular fact.

So why would the roulette wheel spelling out pi be surprising, while having it spell out what it *did* spell out not be surprising? I think the answer lies in the fact that we expect events to be as generic as possible – we expect them to have properties that could reasonably be anticipated in advance, but we expect them not to have any other noteworthy properties. So, for instance, we expect the number of red and black outcomes in the running of the roulette wheel to be roughly equal, and we expect the list of outcomes *not* to have any noteworthy numerical properties other than those dictated by probability theory. Because of this, it is surprising if the roulette wheel spells out pi, but unsurprising if it spells out some otherwise nondescript sequence.

One source of dissatisfaction with what I have just said is that I use the phrase 'noteworthy properites'. Which properties are 'noteworthy', and on what basis do we draw a distinction between noteworthy and unnoteworthy properties? I think the answer to such questions must involve psychological facts about the scientific and mathematical temperaments of the intellectual community. *We* would be stunned if some quantity with no apparent relation to pi turned out to be equal to pi – so much so that we would probably devote great resources to finding out whether the quantity *really was* related to pi after all. But an intellectual community in another age, composed of people with different prejudices and preferences, would perhaps not be so struck. The claims about 'surprise' that I offer here are manifestly *subjective*, insofar as I do not claim to be able to make sense of the idea of something being surprising independently of our expectations, prejudices, and preferences.

Applying these claims once more, we see that there is nothing surprising about $c = 3.00 \times 10^8\ m/s$, even though there would be something surprising about c spelling out the digits of pi. This is because we expect c to have those numerical properties dictated by theory, but no other noteworthy mathematical properties.

Most importantly, I think these considerations show that it *is* appropriate to be surprised at the fact that the language of nature is mathematics (in either the sense described in

Chapter 6 or this chapter.) Given our expectation that the world be generic in its features – i.e., that it have no noteworthy features other than those which could be reasonably anticipated in advance – it would be surprising for the language of nature to be mathematics (in either sense). Thus, Wigner's sense of surprise at the ideal descriptive problem can be justified, even though (i) the use of stronger adjectives seems unwarranted, and (ii) he is mistaken at being surprised at any other version of the problem of the unreasonable effectiveness of mathematics in physics that I have identified.

APPENDIX A

FEYNMAN'S PATH INTEGRAL.

A.1 INTRODUCTION.

In Chapter 3, I discussed the Feynman path integral, claiming that it was a non-rigorous mathematical concept. In this appendix, I would like to discuss the technical details behind this claim.

When Feynman first introduced the path integral into quantum mechanics in the 1940s, he was criticized for trafficking in an idea that was both crazy, and entirely lacking in mathematical rigor.[1] On the charge of craziness, Feynman's critics have been proven wrong; the path integral has turned out to be one of the most important concepts of late 20th century physics. Not only has it changed the face of quantum mechanics, it has also found applications in areas of physics well outside quantum mechanics, and even in areas of mathematics seemingly distant from physics.

On the charge of lacking mathematical rigor, however, Feynman's critics were not so badly wrong. Mathematicians have struggled for a long time to try and make Feynman's path integral rigorous. While they have enjoyed some success (see, for instance, [65] and [49]) such successes have seemed to come with strings attached – they apply only to a limited class of potentials, they do not generalize to the case of field theory, or they have some other unattractive blemish. Even where limited success has been achieved, the idea that the Feynman integral should be an 'integral' in the usual measure theoretic sense has all but

[1] See, for example, section 2.4 of [61].

been abandoned, even though many heuristic arguments involving the path integral explicitly draw on measure theoretic intuitions. In many argumentative contexts, the Feynman integral therefore remains ill-defined.

One finds symptoms of the ill-defined nature of the Feynman integral all over the place. For instance, in Chapter 6 of [77], Rivers argues that the set of paths of finite action (i.e., $S < \infty$) has measure 0. On the other hand, in Section 2.3 of [93], Swanson argues that paths with infinite action do not contribute to the path integral. Both of these claims cannot be true! Yet both Rivers and Swanson give perfectly reasonable heuristic arguments to support their conclusions. (Rivers' argument involves a Fourier transformation of the path integral, while Swanson's argument involves the Riemann-Lebesgue Lemma.) Such disagreements are to be expected when dealing with an ill-defined concept such as the path integral. Indeed, one might even wonder whether there is any fact of the matter about which of Rivers and Swanson is right.

There are all sorts of philosophical and foundational questions that such stories raise. What should our attitude to ill-defined concepts in physics be? How is it that a science as brutally mathematical as physics can survive, and even flourish, with ill-defined concepts? How do physicists insulate themselves from disaster when dealing with such concepts? (Note, after all, that although Rivers and Swanson disagree about subtle issues surrounding paths of infinite action, they do not seem to disagree about the transition amplitudes for any particular quantum system. Some sort of insulation mechanism is surely in place here.) What sort of attitude should we take to the project of trying to make a concept like the Feynman integral rigorous? Must such projects be completed before we are to say that any of Feynman's claims in [32] count as *knowledge*? Would it matter if the mathematicians, after a long time, enjoyed no success in trying to completely rigorize such concepts? What should we say about the epistemological status of Feynman's original arguments? (Of particular interest is Feynman's highly impressionistic argument in section 2.3 of [32] in which he claims that his path integral formalism yields the classical limit whenever the action $S \gg \hbar$.) Given that modern attempts to rigorize the path integral are generally not measure theoretic in nature, what are we to say about the arguments of Feynman that seem quite explicitly to revolve around the idea that expressions of the form $exp\,[\frac{i}{\hbar}S[x(t)]]$ are being *added* when a path integral is evaluated?

I obviously cannot treat all of these questions here.[2] In the present paper, I simply wish to examine the reasons why a simple measure theoretic interpretation of the path integral is impossible. The first mention of difficulties with a naive measure theoretic interpretation of the path integral is to be found in Cameron's 1960 article [15]. Cameron's discussion of these difficulties is brief, however, and applies to only one attempt to try and rigorously define a path integral measure. One is left wondering whether more imaginative attempts might be more successful. The 'no-go' theorem presented in section 3.1 and the discussion on p. 102 of Johnson and Lapidus [49] suffers from a similar defect, insofar as it makes rather strong assumptions on the properties that a path integral measure must satisfy.[3] [4] Little time has been spent proving satisfying 'no-go' theorems that demolish large classes of attempts to make the path integral rigorous. This is surprising given the amount of energy that has gone into trying to develop non-measure-theoretic approaches to the path integral. The purpose of the bulk of this paper (§§$3-5$) is to try and provide a critique of measure theoretic approaches to the path integral that does not leave the reader with this same sense of dissatisfaction.

In §2, I present an introduction to the path integral. §§$3-5$ consist of a discussion of the mathematical difficulties with a measure-theoretic interpretation of the path integral. (The results there are, as far as I know, new.) I will conclude with a brief discussion in §6 as to what these results teach us, and how the methodology of physics is affected by them.

---

[2]Some of these issues have been discussed in the earlier chapter on mathematical rigor.

[3]Specifically, in Theorem 3.1.3 of [49], Johnson and Lapidus demand that a path measure μ be 'translation invariant' in the sense that, if $f : [0,T] \to \mathbb{R}$ and $g : [0,T] \to \mathbb{R}$ are continuous functions such that $f(0) = g(0) = 0$, then

$$\mu(\{x(t) : [0,T] \to \mathbb{R} \text{ such that } x \text{ is continuous, } x(0) = 0 \text{ and } sup(x(t) - f(t)) < \alpha\})$$

and

$$\mu(\{x(t) : [0,T] \to \mathbb{R} \text{ such that } x \text{ is continuous, } x(0) = 0 \text{ and } sup(x(t) - g(t)) < \alpha\})$$

should be equal. While mathematically appealing, 'translation invariance' of this sort hardly seems physically necessary. In fairness, Johnson and Lapidus are simply trying to capture the intuition that 'all paths be weighed equally'. But one is still left wondering how important such an intuition really is.

[4]Section 4.6 of Johnson and Lapidus' [49], in which *'the nonexistence of Feynman's "measure"'* is demonstrated, appears to largely be a re-iteration of the arguments of [15].

A.2 FEYNMAN'S PATH INTEGRAL.

Imagine a system consisting of single particle moving under the influence of a potential V in one dimension. Let $\mathcal{L}$ be the Lagrangian of the system. The path of the particle over an interval of time $[t_a, t_b]$ can be represented by a function $x(t) : [t_a, t_b] \to \mathbb{R}$. Associated with any such path $x(t)$ is an 'action' $S[x(t)]$, given by the formula

$$S[x(t)] = \int_{t_a}^{t_b} \mathcal{L} dt.$$

The classical path then turns out to be a stationary point for the action functional S. Most often, this simply means that the classical path is the one for which S is a minimum.

Now let us consider quantum mechanics. Imagine that we have found a particle at position x_a at time t_a, and we want to calculate the probability amplitude associated with this particle being found subsequently at position x_b at time $t_b > t_a$. Denote this probability amplitude by $K(b, a)$, where $a = (t_a, x_a)$ and $b = (t_b, x_b)$. Feynman's idea was the following: associate with each path connecting a and b (i.e, with each function $x(t) : [t_a, t_b] \to \mathbb{R}$ satisfying $x(t_a) = t_a$ and $x(t_b) = t_b$) the complex number $exp\,[\frac{i}{\hbar} S[x(t)]]$. Then $K(b, a)$ is just the sum of all these complex numbers over all possible paths. In other words,

$$K(b, a) \quad = \sum_{\substack{x(t):[t_a,t_b]\to\mathbb{R} \\ x(t_a)=x_a, x(t_b)=x_b}} exp\,[\frac{i}{\hbar} S[x(t)]].$$

This idea was introduced by Feynman in his doctoral dissertation, and made famous in his book [32].

How are we to we add up quantities associated with an infinite numbers of paths? What does the $\sum$ mean in the above formula? Obviously, it is something closer to an integral than a sum, and so we might try and write

$$K(b, a) \quad = \int_{\substack{x(t):[t_a,t_b]\to\mathbb{R} \\ x(t_a)=x_a, x(t_b)=x_b}} exp\,[\frac{i}{\hbar} S[x(t)]]\, d\mu,$$

or, adopting notation I shall use henceforth,

$$K(b, a) \quad = \int_{\substack{x(t_b)=x_b \\ x(t_a)=x_a}} exp\,[\frac{i}{\hbar} S[x(t)]]\, d\mu$$

where μ is some 'path measure' over the space of all paths connecting a and b. Our question now becomes: how is the measure μ to be defined?[5]

In [32], Feynman tackles this problem by restricting his attention to polygonal paths of a certain sort. Specifically, for $N \in \mathbb{N}$, let F_N be the set of continuous functions $x(t) : [t_a, t_b] \to \mathbb{R}$ such that $x(t_a) = x_a$, $x(t_b) = x_b$, and x is linear on the intervals $[t_a, t_a + \varepsilon], [t_a + \varepsilon, t_a + 2\varepsilon], ...[t_b - \varepsilon, t_b]$, where $\varepsilon = (t_b - t_a)/N$. Restricting attention to functions in F_N, one can then treat

$$\int_{\substack{x(t_b)=x_b \\ x(t_a)=x_a}}^{x \in F_N} exp\,[\frac{i}{\hbar} S[x(t)]]\; d\mu$$

as an ordinary $(N-1)$-dimensional Lebesgue integral. One then lets $N \to \infty$, and hopes for convergence. (See section 2.4 of [32] for more details.) In the case of the free particle, for instance, we get the equation

$$K(b,a) = \lim_{N \to \infty} \int \int ... \int\; exp\,[\frac{im}{2\hbar\epsilon} \sum_{i=1}^{N} (x_i - x_{i-1})^2]\; (2\pi i\hbar\varepsilon/m)^{-N/2}\;\; dx_1 \,...\, dx_{N-1} \qquad \text{(A.1)}$$

(The term $(2\pi i\hbar\varepsilon/m)^{-N/2}$ is a 'normalization constant' that I will not discuss.) This expression not only converges, but gives the correct quantum mechanical amplitudes for a free particle. (See Chapter 3 of [32].)

The problem, however, is that there are cases in which the analogue of formula (A.1) does not converge as $N \to \infty$, or in which it has several limit points. (This is the topic of

[5] One might think that the question of how such a measure should be defined is of limited interest, insofar as this particular way of calculating the kernel $K(b,a)$ rests on the idea of particles moving through *configuration space*, which does not seem to be a sensible way to think in quantum mechanics. (Furthermore, insofar as our ultimate goal is to represent transition amplitudes between *states* (i.e., wavefunctions), one might think that it is quantities like

$$K(\psi_b, \psi_a) \;\; = \int_{\substack{x(t_b)=x_b \\ x(t_a)=x_a}} \int \int \psi_b(x_b) exp\,[\frac{i}{\hbar} S[x(t)]] \psi_a^*(x_a)\; dx_a dx_b d\mu$$

in which we should be interested.) One should note, however, that path integrals are usually evaluated by numerical methods, and that such methods generally involve selecting a large (but finite) set of paths *through configuration space* with the desired endpoints, and adding up all the values of $exp\,[\frac{i}{\hbar} S[x(t)]]$ to obtain an approximation to

$$K(b,a) \;\; = \int_{\substack{x(t):[t_a,t_b]\to\mathbb{R} \\ x(t_a)=x_a, x(t_b)=x_b}} exp\,[\frac{i}{\hbar} S[x(t)]]\; d\mu.$$

So in spite of any ideological quibbles one might have with letting configuration space into the heart of quantum mechanics, the idea that one is summing over paths *though configuration space* seems central to the idea of a path integral, and is therefore worthy of investigation.

§3.) The polygonal path approximation, therefore, does not define a path measure adequate for dealing with more complicated systems. In light of this, one might wonder whether the polygonal path approach is simply too crass. Countering this intuition, I will prove theorems in §4 and §5 showing that, given certain mathematical requirements, there is *no* path measure μ that renders an unambiguous integration theory over paths, and at the same time gives us the quantum mechanical amplitudes we expect.

Let me introduce one final pair of equations. In section 2.5 of [32], Feynman argues that if we want to calculate the amplitude $K(b,a)$, it suffices to fix an intermediate time $t_a < t_c < t_b$, then add up the amplitudes for the particle to go from x_a at time t_a to x_b at time t_b via q at time t_c. Mathematically,

$$K(b,a) = \int dq\, K(b,c)K(c,a),$$

where $c = (t_c, q)$. Writing this out in full, we have

$$\begin{aligned} &\int_{\substack{x(t_a)=x_a \\ x(t_b)=x_b}} dx(t)\, exp\,[\frac{i}{\hbar}S[x(t)]] \;= \\ &\quad \int dq \int_{\substack{x_1(t_a)=x_a \\ x_1(s)=q}} dx_1(t)\, exp\,[\frac{i}{\hbar}S[x_1(t)]] \int_{\substack{x_2(s)=q \\ x_2((t_b)=x_b}} dx_2(t)\, exp\,[\frac{i}{\hbar}S[x_2(t)]] \end{aligned} \tag{A.2}$$

I shall call (A.2) *Feynman's factorization identity.*

This identity can be generalized. Let F_{t_a,t_b} be a functional acting on $\{x(t) : t_a \leq t \leq t_b\}$. Fix $s \in (t_a, t_b)$. For any function $x(t)$ on $[t_a, t_b]$, let $x_1(t)$ be the restriction of $x(t)$ to $[t_a, s]$, and let $x_2(t)$ be the restriction of $x(t)$ to $[s, t_b]$. Assume that there exist functionals $F_{t_a,s}$ acting on $\{x(t) : t_a \leq t \leq s\}$ and F_{s,t_b} acting on $\{x(t) : s \leq t \leq t_b\}$ such that, for all functions $x(t)$ defined on $[t_a, t_b]$, $F_{t_a,t_b}[x(t)] = F_{t_a,s}[x_1(t)]F_{s,t_b}[x_2(t)]$. We then say that F_{t_a,t_b} is *factorizable* into $F_{t_a,s}$ and F_{s,t_b}, written $F_{t_a,t_b} = F_{t_a,s} \times F_{s,t_b}$. We generalize (A.2) by demanding that:

$$\begin{aligned} &\int_{\substack{x(t_a)=x_a \\ x(t_b)=x_b}} dx(t)\, F_{t_a,t_b}[x(t)] \;= \\ &\quad \int dq \int_{\substack{x_1(t_a)=x_a \\ x_1(s)=q}} dx_1(t)\, F_{t_a,s}[x_1(t)] \int_{\substack{x_2(s)=q \\ x_2((t_b)=x_b}} dx_2(t)\, F_{s,t_b}[x_2(t)] \end{aligned} \tag{A.3}$$

whenever $F_{t_a,t_b} = F_{t_a,s} \times F_{s,t_b}$. Call (A.3) the *generalized factorization identity.*

Note that $exp\,[\frac{i}{\hbar}S[x(t)]$, acting on $\{x(t) : t_a \leq t \leq t_b\}$ is factorizable in this way for any $s \in (t_a, t_b)$, because

$$exp\,[\frac{i}{\hbar}\int_{t_a}^{t_b} \mathcal{L}dt] = exp\,[\frac{i}{\hbar}\int_{t_a}^{s} \mathcal{L}dt] exp\,[\frac{i}{\hbar}\int_{s}^{t_b} \mathcal{L}dt].$$

Feynman's factorization identity may therefore be viewed as a special case of the generalized factorization identity. [6]

A.3 A FIRST TRY.

In order to make Feynman's path integral formalism rigorous, the first thing we might try to do is rewrite formula (A.1) in the form:

$$K(a,b) = \int_{\substack{x(t_a)=x_a \\ x(t_b)=x_b}} exp\,[\frac{i}{\hbar}\int_{t_a}^{t_b} \mathcal{L}_{free}\,dt]\,d\mu \tag{A.4}$$

where μ is some well defined measure to be constructed over the space of all possible paths $x(t) : [t_a, t_b] \rightarrow \mathbb{R}$ satisfying $x(t_a) = x_a$ and $x(t_b) = x_b$.

How might we define μ? Before defining a measure, we must first define what our measurable sets are to be. Let $t_a < s_1 \leq s_2 < t_b$, and let $\alpha < \beta$ be real numbers. Let $E_{s_1,s_2,\alpha,\beta}$ be the set of functions $x(t)$ satisfying $x(t_a) = x_a, x(t_b) = x_b$ and $\alpha \leq x(s) \leq \beta$ for all $s \in [s_1, s_2]$. Let $\mathcal{E}$ be the smallest $\sigma-$algebra containing all the $E_{s_1,s_2,\alpha,\beta}$. We demand that our measurable sets include all of $\mathcal{E}$.

Trying to take Feynman's formula (A.1) very literally, we might first try to define a measure μ satisfying

$$\lim_{N\rightarrow\infty} \int\int ... \int (2\pi i\hbar\varepsilon/m)^{-N/2}\, dx_1...dx_{N-1} = \int d\mu \tag{A.5}$$

[6] The only instances of (A.3) we actually need are of the form $F = \chi_T[x(t)]\,exp\,[\frac{i}{\hbar}S[x(t)]]$, where χ_T is the characteristic functional for some set of paths T. Specifically, in Theorem 1 we use (A.3) with $F = 1$, which may be viewed as $F = exp\,[\frac{i}{\hbar}S[x(t)]$ with $\mathcal{L} = 0$, and in Theorem 2 we use what amounts to (A.3) with $F = \chi_T[x(t)]\,exp[\frac{i}{\hbar}S[x(t)]]$, where T is a set of paths of the form $\{x(t) : [t_a, t_b] \rightarrow \mathbb{R}$ such that $\alpha < x(s) < \beta\}$, with $\alpha < \beta$ and $t_a < s < t_b$ all fixed.

where these integrals are thought of as ranging over the possible (classical) positions of the particle at times $t_a+\varepsilon$, $t_a+2\varepsilon$, ... $t_b-\varepsilon$, where $\varepsilon=(t_b-t_a)/N$.

Imagine that for countably many $t\in(t_a,t_b)$, there are associated a Borel set of reals $B(t)$. (The Borel sets of reals are those contained in the smallest σ-algebra containing the open sets.) Define $B(t_a)=\{x_a\}, B(t_b)=\{x_b\}$, and for all other $t\in(t_a,t_b)$ for which $B(t)$ is not already defined, let $B(t)=\mathbb{R}$. Let $S_B=\{x(t): t\in[t_a,t_b]\rightarrow x(t)\in B(t)\}$. Then it is easy to see that each $S_B\in\mathcal{E}$. Inspired by (A.5), we write

$$\mu(S_B)=\lim_{N\rightarrow\infty}\int_{B(t_a+\varepsilon)}\int_{B(t_a+2\varepsilon)}...\int_{B(t_b-\varepsilon)}(2\pi i\hbar\varepsilon/m)^{-N/2}\,dx_1...dx_{N-1}$$

where again $\varepsilon=(t_b-t_a)/N$; i.e.,

$$\mu(S_B)=\lim_{N\rightarrow\infty}[\prod_{i=0}^{i=N-1}\mu_l(B(t_a+i\varepsilon))][(2\pi i\hbar\varepsilon/m)^{-N/2}], \tag{A.6}$$

where μ_l is Lebesgue measure.

The problem, however, is that this limit is not uniquely defined. To see why, let $B(t)=\{0\}$ whenever $(t-t_a)/(t_b-t_a)$ is a rational number with a denominator that is a power of 2 (when expressed in simplest form), let $B(t_a)=\{x_a\}, B(t_b)=\{x_b\}$, and let $B(t)=\mathbb{R}$ for all other $t\in(t_a,t_b)$. Then, if N ranges only over powers of two, the right hand side of (A.6) $\rightarrow 0$ as $N\rightarrow\infty$, but if N ranges only over powers of three, the right hand side of (A.6) $\rightarrow\infty$ as $N\rightarrow\infty$. Which limit should we pick; not just in this case, but quite generally?

Let us consider two alternatives. First, we could define

$$\mu_{inf}(S_B)=\liminf_{N\rightarrow\infty}[\prod_{i=0}^{i=N-1}\mu_l(B(t_a+i\varepsilon))][(2\pi i\hbar\varepsilon/m)^{-N/2}].$$

A second option is

$$\mu_{sup}(S_B)=\limsup_{N\rightarrow\infty}[\prod_{i=0}^{i=N-1}\mu_l(B(t_a+i\varepsilon))][(2\pi i\hbar\varepsilon/m)^{-N/2}].$$

The problem, however, is that neither μ_{inf} nor μ_{sup} are additive, and so neither are measures.

In order to see why, I shall first need to construct sets C and D such that $0 < \mu_{inf}(C) < \infty$ and $0 < \mu_{sup}(D) < \infty$. Assume $(t - t_a)/(t_b - t_a)$ is a rational number that, expressed in simplest form, is q/p where p is some prime. Then let

$$B_1(t) = [\,0, \frac{2\pi i\hbar(t_b - t_a)}{mp}^{p/2(p-1)}].$$

As before, let $B_1(t_a) = \{x_a\}, B_1(t_b) = \{x_b\}$, and for all other $t \in (t_a, t_b)$ for which $B_1(t)$ is not yet defined, let $B_1(t) = \mathbb{R}$. Let $C = S_{B_1}$. If $N = p$ is prime, then

$$\begin{aligned}\prod_{i=0}^{i=N-1} & \mu_l(B_1(t_a + i\varepsilon))[(2\pi i\hbar\varepsilon/m)^{-N/2}] \\ &= ((2\pi i\hbar(t_b - t_a)/mp)^{(p/2(p-1))})^{p-1}.(2\pi i\hbar\varepsilon/m)^{-p/2} \\ &= 1.\end{aligned}$$

If N is not prime, then

$$\prod_{i=0}^{i=N-1} \mu_l(B_1(t_a + i\varepsilon))[(2\pi i\hbar\varepsilon/m)^{-N/2}] = \infty.$$

Thus $\mu_{inf}(C) = 1$.

Let us now construct D. If $(t - t_a)/(t_b - t_a)$ is a rational number of the form q/p, where p is some prime, let $B_2(t) = B_1(t)$. If $(t - t_a)/(t_b - t_a)$ is a rational number *not* of this form, let $B_2(t) = \{0\}$. Again, let $B_2(t_a) = \{x_a\}, B_2(t_b) = \{x_b\}$, and for all other $t \in (t_a, t_b)$, let $B_2(t) = \mathbb{R}$. Let $D = S_{B_2}$. As before, if N is prime,

$$\prod_{i=0}^{i=N-1} \mu_l(B_2(t_a + i\varepsilon))[(2\pi i\hbar\varepsilon/m)^{-N/2}] = 1.$$

For all other N,

$$\prod_{i=0}^{i=N-1} \mu_l(B_2(t_a + i\varepsilon))[(2\pi i\hbar\varepsilon/m)^{-N/2}] = 0.$$

Thus $\mu_{sup}(D) = 1$.

Now, define $K_1(t)$, $K_2(t)$ and $K_3(t)$ as follows: if $(t - t_a)/(t_b - t_a) = 1/4$, let $K_1(t) = [0, 1], K_2(t) = (1, 2]$ and $K_3(t) = [0, 2]$. For all other values of t, let $K_1(t) = K_2(t) = K_3(t) = B_1(t)$. Then it is easily seen that $\mu_{inf}(S_{K_1}) = \mu_{inf}(S_{K_2}) = \mu_{inf}(S_{K_3}) = \mu_{inf}(C) = 1$. But

$S_{K_3} = S_{K_1} \cup S_{K_2}$ and $S_{K_1} \cap S_{K_2} = \emptyset$, and so we should have $\mu_{inf}(S_{K_3}) = \mu_{inf}(S_{K_1}) + \mu_{inf}(S_{K_2})$. Thus μ_{inf} is not additive, is therefore not a measure.

Let us turn our attention to μ_{sup}. Define K_1, K_2 and K_3 anew as follows: if $(t - t_a)/(t_b - t_a) = 1/4$, then let $K_1(t), K_2(t)$ and $K_3(t)$ be as before. Otherwise, let $K_1(t) = K_2(t) = K_3(t) = B_2(t)$. Again, it is easily seen that $\mu_{sup}(S_{K_1}) = \mu_{sup}(S_{K_2}) = \mu_{sup}(S_{K_3}) = \mu_{sup}(D) = 1$. But again, $S_{K_3} = S_{K_1} \cup S_{K_2}$ and $S_{K_1} \cap S_{K_2} = \emptyset$, and so we should have $\mu_{sup}(S_{K_3}) = \mu_{sup}(S_{K_1}) + \mu_{sup}(S_{K_2})$. So μ_{sup} is also not additive, is therefore also not a measure.

We must conclude that, as it stands, there is little sense to be made of formula (A.5). While Feynman's 'polygonal path approximation' works in the simplest case of the free particle, it does not provide us with an unambiguously defined measure that is capable of handling more complex situations.[7]

A.4 A SECOND TRY.

Let us forget about formula (A.5) and polygonal approximations quite generally, and try to construct a path measure μ ourselves, in hope that we will at least be able to take formula (A.4) literally. (This, after all, is all that really matters.)

First note that we must not just construct a single path integral measure, but a whole family of path integral measures. This is because if we want to use the factorization identities (A.2) or (A.3), we not only need a measure μ over $\{x(t) : [t_a, t_b] \to \mathbb{R}$ with $x(t_a) = x_a, x(t_b) = x_b\}$, but we also need measures μ_1 and μ_2 over $\{x(t) : [t_a, s] \to \mathbb{R}$ with $x(t_a) = x_a, x(s) = q\}$ and $\{x(t) : [s, t_b] \to \mathbb{R}$ with $x(s) = q, x(t_b) = x_b\}$ for arbitrary s and q. We therefore need to construct a family of measures $\{\mu_{(t_a,x_a,t_b,x_b)}\}$ for arbitrary t_a, x_a, t_b and x_b. (If desired, we can restrict t_a and t_b to lie in some predetermined range $[T_1, T_2]$, and not bother defining any measures involving times outside of $[T_1, T_2]$. Nothing that follows depends on whether

[7]One might object that sets such as C, D, S_{K_1}, S_{K_2} and S_{K_3} are 'unphysical', and are therefore not the sorts of sets whose measure we should care about. This attitude is misguided – insofar as such sets are contained in the smallest σ-algebra containing all the $E_{s_1,s_2,\alpha,\beta}$, it must be possible to assign them a measure. Unless we assign a measure to all such sets, it will not be possible to develop a useful integration theory.

this is done.) We will often write $\int ... d\mu$ rather than the more cumbersome $\int ... d\mu_{t_a,x_a,t_b,x_b}$, however, when the context makes it clear what is meant.

We say that a family of measures $\{\mu_{(t_a,x_a,t_b,x_b)}\}$ has the *factorization property* if the factorization identity (A.3) holds for any $\mu_{(t_a,x_a,t_b,x_b)}$-measurable functional F that, for some s, can be written as $F_1 \times F_2$, where F_1 and F_2 are $\mu_{(t_a,x_a,s,q)}$-measurable, and $\mu_{(s,q,t_b,x_b)}$-measurable respectively, for all q.

Let us briefly remind ourselves of the definition of an integral over a measure space. Let f be a measurable function which is nonnegative everywhere (i.e., $f \geq 0$). If E is a measurable set, define

$$\int_E f d\mu = \sup \sum_j [\inf_{x \in E_j} f(x)]\, \mu(E_j) \tag{A.7}$$

where the supremum is taken over all partitions $E = \bigcup_j E_j$ of E into a *finite* number of disjoint measurable sets.

For general measurable $f : \mathbb{R} \to \mathbb{R}$, write $f = f^+ - f^-$, where $f^+ = max\{f, 0\}$ and $f^- = -min\{f, 0\}$. Then $f^+ \geq 0$, and $f^- \geq 0$, so we may define

$$\int_E f d\mu = \int_E f^+ d\mu \; - \int_E f^- d\mu.$$

(Note that at least one of $\int_E f^+ d\mu$ and $\int_E f^- d\mu$ must be finite for this definition to make sense.)

Finally, for measurable $f : \mathbb{R} \to \mathbb{C}$, write $f = u + iv$, where u and v are real valued, and define

$$\int_E f d\mu = \int_E u d\mu \; + \; i \int_E v d\mu.$$

For arbitrary t_a, x_a, t_b and x_b, define

$$V_{(t_a,x_a,t_b,x_b)} = \{x(t) : x(t_a) = x_a \text{ and } x(t_b) = x_b\}$$

Also, define

$$c_{(t_a,x_a,t_b,x_b)} = \mu_{(t_a,x_a,t_b,x_b)}(V_{(t_a,x_a,t_b,x_b)}).$$

Then we have the following lemma:

Lemma. *Assume μ is a measure such that, for some t_a, x_a, t_b, x_b and $\mathcal{L}$,*

$$\int_{\substack{x(t_a)=x_a \\ x(t_b)=x_b}} exp\,[\frac{i}{\hbar}\int_{t_a}^{t_b} \mathcal{L}\,dt]\,d\mu$$

is defined and finite. Then $c_{(t_a,x_a,t_b,x_b)}$ is finite.

Proof. If

$$\int_{\substack{x(t_a)=x_a \\ x(t_b)=x_b}} exp\,[\frac{i}{\hbar}\int_{t_a}^{t_b} \mathcal{L}\,dt]\,d\mu$$

is defined and finite, then for any $k \in \mathbb{R}$,

$$\int_{\substack{x(t_a)=x_a \\ x(t_b)=x_b}} exp\,[\frac{i}{\hbar}\int_{t_a}^{t_b} \mathcal{L}'\,dt]\,d\mu$$

is defined and finite, where $\mathcal{L}' = \mathcal{L} + k/(t_b - t_a)$. Consequently, its real part,

$$\int_{\substack{x(t_a)=x_a \\ x(t_b)=x_b}} cos\,[\frac{1}{\hbar}(S[x(t)] + k)]\,d\mu$$

must also be finite for any k, where we have written $S[x(t)] = \int_{t_a}^{t_b} \mathcal{L}\,dt$. From this it follows that

$$\int_{\substack{x(t_a)=x_a \\ x(t_b)=x_b}} cos^{+}\,[\frac{1}{\hbar}(S[x(t)]] + k)]\,d\mu$$

must be finite, where for all θ, $cos^{+}(\theta) = \max(cos(\theta), 0)$. For $0 < \varepsilon < 1$, let

$$E_{\varepsilon,k} = \{x(t) : cos^{+}\,[\frac{1}{\hbar}(S[x(t)] + k)] > \varepsilon\},$$

and let

$$F_{\varepsilon,k} = \{x(t) : cos^{+}\,[\frac{1}{\hbar}(S[x(t)] + k)] \leq \varepsilon\}.$$

Then, from (A.7),

$$\int_{\substack{x(t_a)=x_a \\ x(t_b)=x_b}} cos^{+}\,[\frac{1}{\hbar}(S[x(t)]] + k)]\,d\mu \geq$$
$$[\inf_{x(t)\in E_{\varepsilon,k}} cos^{+}\,(\tfrac{1}{\hbar}(S[x(t)]] + k))\,]\mu(E_{\varepsilon,k}) + [\inf_{x(t)\in F_{\varepsilon,k}} cos^{+}\,(\tfrac{1}{\hbar}(S[x(t)]] + k))\,]\mu(F_{\varepsilon,k})$$

Because $[\inf_{x(t)\in E_{\varepsilon,k}} cos^{+}\,(\frac{1}{\hbar}(S[x(t)]] + k))\,] > 0$, it follows that we must have $\mu(E_{\varepsilon,k}) < \infty$, or else

$$\int_{\substack{x(t_a)=x_a \\ x(t_b)=x_b}} cos^{+}\,[\frac{1}{\hbar}(S[x(t)]] + k)]\,d\mu$$

would be infinite. So for all k, $\mu(E_{\varepsilon,k}) < \infty$. However, for any ε, there exists a finite list of reals $k_1, k_2, ...k_r$ such that, for any $x(t)$, $x(t) \in E_{\varepsilon,k_i}$ for some $i \in \{1, 2, ..., r\}$. (This follows from elementary properties of the cosine function.) Thus $V_{(t_a,x_a,t_b,x_b)}$ is the union of finitely many sets of finite measure, and thus $c_{(t_a,x_a,t_b,x_b)}$ is finite. □

Let me introduce one further technicality before stating and proving the main theorem of this section. In the case of a free particle, we know in advance what

$$K(b,a) = \int_{\substack{x(t_a)=x_a \\ x(t_b)=x_b}} exp\,[\frac{i}{\hbar}\int_{t_a}^{t_b} \mathcal{L}_{free}\, dt]\, d\mu$$

should work out to be; specifically, we should have

$$K(b,a) = \sqrt{\frac{m}{2\pi i\hbar(t_b - t_a)}}\; exp\,[\frac{im(x_b - x_a)^2}{2\hbar(t_b - t_a)}].$$

(See, for example, section 3.1 of [93].) Say that a family of measures $\{\mu_{(t_a,x_a,t_b,x_b)}\}$ reproduces the quantum amplitudes for a free particle just in case, for all t_a, x_a, t_b and x_b,

$$\int_{\substack{x(t_a)=x_a \\ x(t_b)=x_b}} exp\,[\frac{i}{\hbar}\int_{t_a}^{t_b} \mathcal{L}_{free}\, dt]\, d\mu = \sqrt{\frac{m}{2\pi i\hbar(t_b - t_a)}}\; exp\,[\frac{im(x_b - x_a)^2}{2\hbar(t_b - t_a)}].$$

We then have the following theorem:

Theorem 3. *There is no family of measures $\{\mu_{(t_a,x_a,t_b,x_b)}\}$ that has the factorization property, and reproduces the quantum amplitudes for a free particle.*

Proof. Assume, for a contradiction, that we have a family $\{\mu_{(t_a,x_a,t_b,x_b)}\}$ of measures having the factorization property, and producing the correct quantum amplitudes for a free particle. Fix t_a, x_a, t_b and x_b; we show that $c_{(t_a,x_a,t_b,x_b)} = \infty$, contradicting the previous lemma.

Recall that the inequality

$$|\int_X f\, d\mu| \leq \int_X |f|\, d\mu \tag{A.8}$$

holds for any measure μ, integrable function f and measurable set X.

Choose any s such that $t_a < s < t_b$. For any $t < t'$, let $I_{t,t'}$ be the functional acting on $\{x(t) : x(t) : [t,t'] \to \mathbb{R}\}$ such that $I_{t,t'}[x(t)] = 1$ for every $x(t)$ in its domain. Then I_{t_a,t_b} factorizes as $I_{t_a,t_b} = I_{t_a,s} \times I_{s,t_b}$. So

$$c_{(t_a,x_a,t_b,x_b)} = \int_{\substack{x(t_a)=x_a \\ x(t_b)=x_b}} I_{t_a,t_b}\, d\mu$$

$$= \int_{-\infty}^{+\infty} dq \int_{\substack{x(t_a)=x_a \\ x(s)=q}} I_{t_a,s}\, d\mu \int_{\substack{x(s)=q \\ x(t_b)=x_b}} I_{s,t_b}\, d\mu$$

(using the factorization property)

$$= \int_{-\infty}^{+\infty} dq \int_{\substack{x(t_a)=x_a \\ x(s)=q}} |exp\,[\frac{i}{\hbar}S[x(t)]]|\, d\mu \int_{\substack{x(s)=q \\ x(t_b)=x_b}} |exp\,[\frac{i}{\hbar}S[x(t)]]|\, d\mu$$

$$\geq \int_{-\infty}^{+\infty} dq\, |\int_{\substack{x(t_a)=x_a \\ x(s)=q}} exp\,[\frac{i}{\hbar}S[x(t)]]|\, d\mu|| \int_{\substack{x(s)=q \\ x(t_b)=x_b}} exp\,[\frac{i}{\hbar}S[x(t)]]|\, d\mu|$$

(using (A.8))

$$= \int_{-\infty}^{+\infty} dq |K((t_a,x_a),(s,q))||K((s,q)),(t_b,x_b))|$$

(using the fact that our measures reproduce the

quantum amplitudes for a free particle)

$$= \int_{-\infty}^{+\infty} dq\, \frac{m}{2\pi\hbar\sqrt{(s-t_a)(t_b-s)}}$$

$$= \infty$$

i.e., $c_{(t_a,x_a,t_b,x_b)} = \infty$. □

This result suggests that attempts far more general than those discussed in the previous section to construct a countably additive path measure are doomed to failure.[8]

A.5 A THIRD TRY.

In order to take expression (A.4) literally, we tried to directly construct a path integral measure μ. We did not have much luck. Let us try another route. Although in (A.4), we

[8] For an important proviso, see the discussion after the proof of Theorem 2 in the next section.

may not be able to take $d\mu$ literally, perhaps we can somehow assign meaning to the entire subexpression

$$exp\,[\frac{i}{\hbar}\int_{t_a}^{t_b} \mathcal{L}_{free}\, dt]\, d\mu.$$

More generally, for each Lagrangian $\mathcal{L}$, perhaps we can construct a measure $\mu_{\mathcal{L}}$ such that, for instance,

$$K(a,b) = \int_{\substack{x(t_a)=x_a \\ x(t_b)=x_b}} d\mu_{\mathcal{L}}.$$

First note that this means that we must now consider *complex* measures. A complex measure μ on a σ-algebra $\mathcal{M}$ is usually defined as a function $\mu : \mathcal{M} \to \mathbb{C}$ such that $\mu(M) = \sum_i \mu(M_i)$, whenever $M \in \mathcal{M}$, $M_i \in \mathcal{M}$ for all i, $\cup M_i = M$ and $M_i \cap M_j = \emptyset$ for all $i \neq j$. For a discussion of complex measures, see, for instance, Chapter 6 of Rudin [79].

This way of trying to make the path integral rigorous is perhaps the most well known. It also fails. This was first noted by Cameron in [15]. Specifically, he noted that the most natural way of constructing $\mu_{\mathcal{L}_{free}}$ yields a measure with 'unbounded variation'. This, in turn, is well known to lead to violations of coutable additivity. The unbounded variation of traditional ways of defining path integral measures had, strangely enough, been overlooked in early work in path integrals. In fact, Cameron cites a paper by no less than Gelfand and Yaglom [37] containing erroneous results, due to the authors not recognizing the unbounded variation of the path integral measure they had constructed.

Let us define some terminology. For any complex measure μ and measurable set E, define

$$|\mu|(E) = \ \sup \sum_{i=1}^{\infty} |\mu(E_i)|,$$

where the supremum is taken over all countable sets $\{E_i\}$ of measurable sets satisfying $\bigcup_i E_i = E$, and $E_i \cap E_j = \emptyset$ for all $i \neq j$. Then one can prove (see chapter 6 of [79]) that $|\mu|(E) < \infty$ for any measurable E. This is what is meant when it is said that a complex measure must have 'bounded variation'.

A simple argument shows that *any* reasonable attempt to construct $\mu_{\mathcal{L}_{free}}$ leads to a measure of unbounded variation. The argument for this is contained in the theorem below. Before we can state and prove this theorem, however, we need some more definitions.

As before, for each Lagrangian $\mathcal{L}$, we shall need to construct not just one measure $\mu_{\mathcal{L}}$, but a whole family of measures $\{\mu_{\mathcal{L},t_a,x_a,t_b,x_b}\}$. This is necessary in order that we be able to use the Feynman factorization formala, which in this case, is just

$$\int_{\substack{x(t_a)=x_a\\ x(t_b)=x_b}} d\mu_{\mathcal{L},t_a,x_a,t_b,x_b} = \int dq \int_{\substack{x_1(t_a)=x_a\\ x_1(s)=q}} d\mu_{\mathcal{L},t_a,x_a,s,q} \int_{\substack{x_2(s)=q\\ x_2((t_b)=x_b}} d\mu_{\mathcal{L},s,q,t_b,x_b}.$$

Again, we generalize this by demanding that

$$\int_{\substack{x(t_a)=x_a\\ x(t_b)=x_b}} F_{t_a,t_b}[x(t)]\, d\mu_{\mathcal{L},t_a,x_a,t_b,x_b} =$$
$$\int dq \int_{\substack{x_1(t_a)=x_a\\ x_1(s)=q}} F_{t_a,s}[x_1(t)]\, d\mu_{\mathcal{L},t_a,x_a,s,q} \int_{\substack{x_2(s)=q\\ x_2((t_b)=x_b}} F_{s,t_b}[x_2(t)]\, d\mu_{\mathcal{L},s,q,t_b,x_b}$$

whenever $F_{t_a,t_b} = F_{t_a,s} \times F_{s,t_b}$. As before, we say that $\{\mu_{\mathcal{L},t_a,x_a,t_b,x_b}\}$ has the *factorization* property if this later identity holds whenever $F_{t_a,t_b} = F_{t_a,s} \times F_{s,t_b}$.

Say also that the family of measures $\mu_{\mathcal{L},t_a,x_a,t_b,x_b}$ *directly* reproduces the quantum amplitudes for a free particle just in case

$$\int_{\substack{x(t_a)=x_a\\ x(t_b)=x_b}} d\mu_{\mathcal{L},t_a,x_a,t_b,x_b} = \sqrt{\frac{m}{2\pi i\hbar(t_b-t_a)}}\, exp\,[\frac{im(x_b-x_a)^2}{2\hbar(t_b-t_a)}].$$

The following theorem is somewhat more straightforward to prove than Theorem 1:

Theorem 4. *There is no family of measures $\{\mu_{(t_a,x_a,t_b,x_b)}\}$ that has the factorization property, and directly reproduces the quantum amplitudes for a free particle.*

Proof. Assume to the contrary that $\{\mu_{(t_a,x_a,t_b,x_b)}\}$ has the factorization property and directly reproduces the quantum amplitudes for a free particle.

For simplicity, let $t_a = 0, t_b = 1$, and $x_a = x_b = 0$. For real numbers $\alpha < \beta$, define

$$F_{\alpha,\beta} = \{x(t) : x(0) = 0, x(1) = 0, \text{ and } \alpha < x(1/2) < \beta\}.$$

Let $I_{\alpha,\beta}$ be the characteristic functional of $F_{\alpha,\beta}$ (i.e., $I_{\alpha,\beta}[x(t)] = 1$ if $x(t) \in F_{\alpha,\beta}$, and $I_{\alpha,\beta}[x(t)] = 0$ otherwise.) Define functionals $I_{1,\alpha,\beta}$ and $I_{2,\alpha,\beta}$ on the set of functions $x(t) : [0,1/2] \to \mathbb{R}$ and $x(t) : [1/2,1] \to \mathbb{R}$ respectively as follows:

$$I_{1,\alpha,\beta}[x(t)] = 1 \text{ if } x(0) = 0 \text{ and } \alpha < x(1/2) < \beta, \text{ and } I_{1,\alpha,\beta}[x(t)] = 0 \text{ otherwise.}$$

$$I_{2,\alpha,\beta}[x(t)] = 1 \text{ if } x(1) = 0 \text{ and } \alpha < x(1/2) < \beta, \text{ and } I_{2,\alpha,\beta}[x(t)] = 0 \text{ otherwise.}$$

Then $I_{\alpha,\beta} = I_{1,\alpha,\beta} \times I_{2,\alpha,\beta}$. Using the factorization property, we then have

$$\begin{aligned}
\mu_{(0,0,1,0)}(F_{\alpha,\beta}) &= \int_{\substack{x(0)=0\\x(1)=0}} I_{\alpha,\beta}\, d\mu \\
&= \int_{-\infty}^{+\infty} dq \int_{\substack{x(0)=0\\x(1/2)=q}} I_{1,\alpha,\beta}\, d\mu \int_{\substack{x(1/2)=q\\x(1)=0}} I_{2,\alpha,\beta}\, d\mu
\end{aligned}$$

If $q \leq \alpha$ or $q \geq \beta$, then

$$\int_{\substack{x(0)=0\\x(1/2)=q}} I_{1,\alpha,\beta}\, d\mu = \int_{\substack{x(1/2)=q\\x(1)=0}} I_{2,\alpha,\beta}\, d\mu = 0,$$

while if $\alpha < q < \beta$, then

$$\int_{\substack{x(0)=0\\x(1/2)=q}} I_{1,\alpha,\beta}\, d\mu = \int_{\substack{x(0)=0\\x(1/2)=q}} d\mu \quad \text{and} \quad \int_{\substack{x(1/2)=q\\x(1)=0}} I_{2,\alpha,\beta}\, d\mu = \int_{\substack{x(1/2)=q\\x(1)=0}} d\mu.$$

Thus

$$\mu_{(0,0,1,0)}(F_{\alpha,\beta}) = \int_{\alpha}^{\beta} dq \int_{\substack{x(0)=0\\x(1/2)=q}} d\mu \int_{\substack{x(1/2)=q\\x(1)=0}} d\mu.$$

If $\{\mu_{(t_a,x_a,t_b,x_b)}\}$ directly reproduces the quantum amplitudes for a free particle, then

$$\int_{\substack{x(0)=0\\x(1/2)=q}} d\mu = \int_{\substack{x(1/2)=q\\x(1)=0}} d\mu = \sqrt{\frac{m}{\pi i \hbar}}\, exp\,[\frac{im}{\hbar} q^2].$$

So

$$\mu_{(0,0,1,0)}(F_{\alpha,\beta}) = \frac{m}{\pi i \hbar} \int_{\alpha}^{\beta} dq\ \exp\,[\frac{2im}{\hbar} q^2].$$

Let $(\alpha_0, \beta_0), (\alpha_1, \beta_1), (\alpha_2, \beta_2), ...$ be a sequence of disjoint subintervals of $\mathbb{R}$. Then

$$\begin{aligned}
|\mu_{(0,0,1,0)}|(V_{(0,0,1,0)}) &\geq \sum_{i=0}^{\infty} |\mu_{(0,0,1,0)}(F_{\alpha_i,\beta_i})| \\
&= \sum_{i=0}^{\infty} \frac{m}{\pi\hbar} | \int_{\alpha_i}^{\beta_i} dq\ \exp\,[\frac{2im}{\hbar} q^2]| \\
&\geq \sum_{i=0}^{\infty} \frac{m}{\pi\hbar} |\, Im(\int_{\alpha_i}^{\beta_i} dq\ \exp\,[\frac{2im}{\hbar} q^2])| \\
&= \sum_{i=0}^{\infty} \frac{m}{\pi\hbar} | \int_{\alpha_i}^{\beta_i} dq\ \sin\,[\frac{2m}{\hbar} q^2]|.
\end{aligned}$$

For $i \in \{0, 1, 2 \ldots\}$, let $\alpha_i = \sqrt{(\hbar i\pi)/2m}$ and $\beta_i = \hbar\sqrt{(\hbar(i+1)\pi)/2m}$. Then, for each i,

$$| \int_{\alpha_i}^{\beta_i} dq \, \sin [\frac{2m}{\hbar} q^2]| = \int_{\alpha_i}^{\beta_i} dq \, | \sin [\frac{2m}{\hbar} q^2]|$$

Thus,

$$\begin{aligned} |\mu_{(0,0,1,0)}|(V_{(0,0,1,0)}) &\geq \sum_{i=0}^{\infty} \frac{m}{\pi\hbar} \int_{\alpha_i}^{\beta_i} dq \, | \sin [\frac{2m}{\hbar} q^2]| \\ &= \int_0^{\infty} dq \, | \sin [\frac{2m}{\hbar} q^2]| \end{aligned}$$

This last integral, however, diverges. Consequently, $|\mu_{(0,0,1,0)}|(V_{(0,0,1,0)}) = \infty$, which is a contradiction.[9] $\square$

A note must be added to both Theorem 1 and Theorem 2. We have assumed throughout that our definitions of integrals are the standard measure theoretic ones. It must be acknowledged, however, that different definitions the integral are possible. For instance, there are measurable functions $f : \mathbb{R} \to \mathbb{R}$ for which the Lebesgue integral is undefined, but for which an improper Riemann integral may be defined. So even for functions $f : \mathbb{R} \to \mathbb{R}$, it is not as if the Lebesgue integral is the final word to be said on the topic of integration. When trying to formalize the Feynman integral, perhaps we ought to turn our attention to a different definition of the integral – perhaps, for instance, one closer to the improper Riemann integral than the Lebesgue integral?

[9] One might wonder why Theorem 1 cannot now be proved as follows: assume $\{\mu_{(t_a,x_a,t_b,x_b)}\}$ has the factorization property and reproduces the quantum amplitudes for a free particle. Then define

$$\mu^*_{(t_a,x_a,t_b,x_b)}(S) = \int_{\substack{x(t_a)=x_a \\ x(t_b)=x_b}} \chi_S \, \exp [\frac{i}{\hbar} S[x(t)]] \, d\mu_{(t_a,x_a,t_b,x_b)},$$

where χ_S is the characteristic functional of S. One might then try to argue that $\mu^*_{(t_a,x_a,t_b,x_b)}$ has the factorization property and directly reproduces the quantum amplitudes for a free particle, contradicting Theorem 2. The problem, however, is that μ^* need not even be a measure. Unless $c_{(t_a,x_a,t_b,x_b)} < \infty$, the convergence of

$$\int_{\substack{x(t_a)=x_a \\ x(t_b)=x_b}} \chi_S \, \exp [\frac{i}{\hbar} S[x(t)]] \, d\mu_{(t_a,x_a,t_b,x_b)}$$

is at best conditional. It follows from this that one will be able to find partitions U_i and V_i of the space of paths such that

$$\lim_{n\to\infty} \sum_{i=0}^{n} \mu^*(U_i) \neq \lim_{n\to\infty} \sum_{i=0}^{n} \mu^*(V_i).$$

Thus, μ^* fails to be a measure.

To add further grist to this mill, recall that one reason that mathematicians often cite for preferring the Lebesgue integral over Riemann-type integrals is that functions like

$$f(x) = 1 \text{ if } x \text{ is rational, and } 0 \text{ otherwise,}$$

are Lebesgue integrable, but not Riemann integrable. But one might surely question whether this sort of generality is really useful in physics. (In fairness, there are other reasons for making the Lebesgue integral the integral of choice; in particular, it allows us to manipulate limits with much greater freedom than in the case of the Riemann integral. To what extent the physicist can abandon this is a tricky question.)

For our purposes, however, it suffices to note that an integral like the improper Riemann integral is not a purely *measure theoretic* integral, insofar as it involves a limiting procedure that must be specified in advance. The main result of this section can therefore be summarized as follows: the formula (A.4) cannot be thought of as an integral *in the usual measure theoretic sense.* Whether or not some other theory of integration can handle the Feynman integral better is a topic beyond the scope of the present paper. (In fact, in special cases, one *can* make sense of the Feynman integral using alternate theories of integration; see, in particular, [49]. It must be emphasized, however, that such methods do not work in general, and are not entirely free of blemishes.)

One final observation: it is interesting to note that while one *cannot* construct a family of path measures that has the factorization property and can be used to solve quantum mechanics problems, one *can* construct a family of path measures that has the factorization property and can be used to solve a different set of physics problems. Consider the heat equation:

$$\frac{\partial}{\partial t}u(x,t) = \frac{1}{2}\frac{\partial^2}{\partial x^2}u(x,t)$$

Ignoring real constants, this is the same as the free Schrödinger equation without the i. One can construct a kernel $K^*(a,b)$ to represent the way heat propagates under this equation; it turns out that we would like it to obey the relation

$$K^*(a,b) = \int_{\substack{x(t_a)=x_a \\ x(t_b)=x_b}} exp\,[-\int_{t_a}^{t_b} \mathcal{L}_{free}\,dt]\,d\mu,$$

where μ is a path measure. (Note the absence of any i in the exponential.) For reasons similar to those discussed in the previous section, no such μ can be constructed. We might then try to construct a measure $\mu_{\mathcal{L}}$ such that

$$K^*(a,b) = \int_{\substack{x(t_a)=x_a \\ x(t_b)=x_b}} d\mu_{\mathcal{L}},$$

i.e., we might instead try to construct a measure $\mu_{\mathcal{L}} = exp\,[-\int_{t_a}^{t_b} \mathcal{L}_{free}\,dt]\,d\mu$. In this case, it *is* possible to rigorously construct such a measure; it is known as the *Wiener measure.* (See chapter 3 of [39], and chapter 3 of [49].) The reader should note that the integral

$$\int_{-\infty}^{\infty} dq\ \exp\,[cq^2]$$

converges *absolutely* for real c, and so we do not run into the countable additivity problems discussed above. We may therefore conclude that it is the i in Schrödinger's equation – absent in the heat equation – that causes all the problems in defining path integrals!

A.6 USING THE PATH INTEGRAL.

Given all the conceptual difficulties in rigorously formulating the path integral, one might wonder how the physicist finds any meaning a formula like

$$K(b,a) \quad = \int_{\substack{x(t_b)=x_b \\ x(t_a)=x_a}} exp\,[\frac{i}{\hbar}S[x(t)]]\,d\mu.$$

How do physicists extract so much information from mathematically questionable formulae like this? One might imagine the answer: they do so *with great care.* But this answer is too glib – after all, why should physicists be able to extract *any* useful information from an ill-defined (and perhaps even mathematically incoherent) concept, regardless of how much 'care' is taken?

I cannot tackle this question in full generality here, but part of the answer lies in the fact that one can often extract useful information from falsehoods, or even incoherent concepts, provided that one restricts the inferential role that such falsehoods or concepts are permitted to play in an argument. A classic example of this is the use of the delta function in elementary quantum mechanics. Provided that one keeps one's delta functions within an integral sign, nothing can go wrong; in fact, many calculations are considerably simplified. But if one interrogates the delta function outside of an integral, one quickly ends up in trouble. Thus, the physicist restricts his inferences involving delta functions to those in which the delta function remains safely tucked away within an integral. The physicist uses the mathematically incoherent delta function in an *inferentially restrictive way* – he does not trust the conclusion of simply *any* mathematically valid argument involving the delta function, but only those conclusions that may be obtained using some specific, restricted set of inferences.

Much the same is true of the path integral. The mathematics of the previous sections shows that the path integral cannot be construed as an integral in the ordinary measure theoretic sense. The physicist copes with this by introducing inferential restrictions that must be obeyed in any argument involving path integrals. One specific inferential restriction is that he avoids arguments involving the path integral measure separated from the context of an actual path integral.

Recall, for instance, the disagreement between Rivers and Swanson discussed in the introduction, and note how it revolves around the question of the measure of the set of paths of infinite action. This is precisely the sort of question that the physicist will generally shy away from. While both Rivers and Swanson take their discussions to have heuristic value, neither base any of their subsequent calculations on their view about the measure of the set of paths of infinite action. This is no coincidence – the careful physicist will make sure that nothing of real importance hangs on arcane details about the measure of specific sets. The physicist knows that to do so would be to ask for trouble, just as to reason about delta functions outside an integral sign is a recipe for disaster.

This commitment to inferential restrictiveness introduces significant new burdens to the physicist. One way in which this is particularly evident is in the treatment of constraints within the path integral formalism. Intuitively, one might think that one could incorporate constraints into the path integral formalism by restricting the set of classical paths over which one integrates the expression $exp\,[\frac{i}{\hbar}S[x(t)]]$. For instance, imagine the 1-dimensional problem of a particle trapped inside an infinite potential well, defined by

$$V(x) = 0 \text{ if } 0 < x < L, \text{ and } V(x) = \infty \text{ otherwise.}$$

In order to calculate $K(b, a)$ for this system, one might define

$$P = \{x(t) : [t_a, t_b] \to [0, L] \text{ such that } x(t_a) = x_a \text{ and } x(t_a) = x_a\},$$

and then try to let

$$K(b, a) = \int_{\substack{x(t_b)=x_b \\ x(t_a)=x_a}}^{P} exp\,[\frac{i}{\hbar} S_{free}[x(t)]]\, d\mu,$$

where $S_{free}[x(t)]$ is the classical action associated with a free particle, and the superscript P on the integral sign means that the integral is only to be evaluated over paths in P. However, integrating over a restricted domain P is dangerously close to asking questions about the measures of paths with a given action, and thus the physicist must be careful not to put too much faith in this expression for $K(b, a)$. (Indeed, it turns out that insofar as this expression for $K(b, a)$ is meaningful, it gives the wrong answer.)

Instead, in order to calculate $K(b, a)$ for a constrained system, one generally figures out a way of expressing $K(b, a)$ in terms of *unconstrained* path-integrals, i.e., path integrals evaluated over *all* paths through some (perhaps different) phase space, subject to some given initial and final conditions. (One way of doing this is to use the method of Lagrange multipliers; see for example, section 4.6 of [93]. There are many cases, however, in which the method of Lagrange multipliers does not work. In such cases, incorporating constraints can be difficult indeed). In the case of the particle in an infinite potential well, one can derive

$$K(t_b, x_b, t_a, x_a) = \sum_{i \in \mathbb{Z}} (K_{free}(t_b, x_b - 2iL, t_a, x_a) - K_{free}(t_b, 2iL - x_b, t_a, x_a))$$

where $K_{free}(b, a)$ is the propagation amplitude for an *unconstrained* free particle, i.e., a free particle not confined to a box.

The general lesson is that physicists must pay a price for inferential restrictiveness, insofar as attractive mathematical forms of argument become unavailable to them, and new forms of argument must be sought. Sometimes, however, the benefits outweigh the costs. In such cases, it can be worth allowing one's methods to fly in the face of mathematical rigor. The path integral is an especially good example of this.

APPENDIX B

THE FORCING RESULT

In this Appendix, I would like to introduce the technique of forcing, and use it to sketch a proof of the Main Theorem discussed in Chapter 6. My discussion of forcing will be largely standard (see references below). The application of forcing found in the Main Theorem is also largely standard, with the exception of a few small details that I shall signal as they occur.

The Main Theorem in question is the following:

Main Theorem: $Con(ZFC) \rightarrow Con(ZFC+$there is a Dedekind cut of $\mathbb{Q}^V$ not contained in $\mathbb{R}^V$.)

The proof of this theorem involves the technique of forcing, developed by Cohen in [20, 21]. The machinery associated with this technique is quite complex, and so I shall only present a few important and standard details. The interested reader should consult [48] or [53] for a thorough and rigorous mathematical treatment. My exposition is largely extracted from Chapter 3 of Jech's [48].

One way of developing the machinery of forcing involves the use of Boolean algebras. I shall begin by presenting some basic definitions:

Definition: A *Boolean Algebra* is a set B with at least two elements, 0 and 1, together with binary operations $\wedge$ and $\vee$ and a unary operation $\neg$ satisfying the following axioms:

(i) $u \wedge v = v \wedge u,\ u \vee v = v \vee u$ (commutativity)

(ii) $u \wedge (v \wedge w) = (u \wedge v) \wedge w,\ u \vee (v \vee w) = (u \vee v) \vee w$ (associativity)

(iii) $u \wedge (v \vee w) = (u \vee v) \wedge (u \vee w),\ u \vee (v \wedge w) = (u \wedge v) \vee (u \wedge w)$ (distributivity)

(iv) $u \wedge u = u,\ u \vee u = u$

(v) $u \wedge (u \vee v) = u,\ u \vee (u \wedge v) = u$

(vi) $u \vee 0 = u,\ u \vee 1 = 1,\ u \wedge 0 = 0,\ u \wedge 1 = u$

(vii) $u \vee \neg u = 1,\ u \wedge \neg u = 0,\ \neg\neg u = u$

(viii) $\neg(u \vee v) = (\neg u) \wedge (\neg v),\ \neg(u \wedge v) = (\neg u) \vee (\neg v)$ (deMorgan's laws)

Definition: If u and v are elements of a Boolean Algebra, define $u \leq v$ iff $u \vee v = v$.

Definition: Let B be a Boolean Algebra. A *filter* on B is a subset I of B such that

(i) $1 \in F,\ 0 \notin F$

(ii) if $u, v \in F$, then $u \wedge v \in F$

(iii) if $u, v \in B$, $u \in F$ and $u \leq v$, then $v \in F$.

An *ultrafilter* F satisfies the additional property that for every $u \in B$, $u \in F$ or $\neg u \in F$.

Definition: Let B be a Boolean Algebra. Say that B is *complete* iff each subset X of B has a greatest lower bound and a least upper bound. If B is complete and X is a nonempty subset of B, define

$$\bigvee\{u : u \in X\} = \text{ the least upper bound of } X,$$

and

$$\bigwedge\{u : u \in X\} = \text{ the greatest lower bound of } X.$$

Also define $\bigvee \emptyset = 0$ and $\bigwedge \emptyset = 1$.

Definition: Let B be a complete Boolean Algebra, and let $\mathcal{M}$ be a set of subsets of B. An ultrafilter G on B is $\mathcal{M}$ *generic* iff $X \subseteq G$ and $X \in \mathcal{M}$ implies $\bigwedge X \in G$.

All the definitions presented thus far belong to the classical theory of Boolean Algebras. We now consider some concepts of a more set-theoretic nature. In what follows, let V denote the set-theoretic universe.

The language of set theory consists of two binary relations, '$=$' and '$\in$'. A *structure* of set theory may therefore be taken to be a triple $<U, I, E>$, where U is a class, and I, E are binary relations corresponding to $=$ and $\in$. Another way of writing this is to say that there are functions $\| x = y \|$ and $\| x \in y \|$ from $U \times U$ to $\{0, 1\}$ such that, for any $x, y \in U$,

(i) $\| x = y \| = 1$ if $x = y$, and $\| x = y \| = 0$ otherwise.

(ii) $\| x \in y \| = 1$ if $x \in y$, and $\| x \in y \| = 0$ otherwise.

We now generalize the notion of a structure of set theory so that the truth values of sentences are no longer confined to the set $\{0, 1\}$, but instead take on arbitrary values from some complete Boolean Algebra B.

Definition: Let B be a complete Boolean Algebra. A *Boolean valued structure* $\mathbf{U}$ *of the language of set theory* consists of a set U and two binary functions $\| x = y \|$ and $\| x \in y \|$ from $U \times U$ to B such that:

(i) $\| x = x \| = 1$

(ii) $\| x = y \| = \| y = x \|$

(iii) $(\| x = y \| \wedge \| y = z \|) \leq \| x = z \|$

(iv) $(\| x \in y \| \wedge \| x = v \| \wedge \| y = w \|) \leq \| v \in w \|$

Given such a structure $\mathbf{U}$, for every formula $\phi(x_1, x_2, ...x_n)$ in the language of set theory, and every $a_1, a_2, ...a_n \in U$, we define the *Boolean value* of $\phi(a_1, a_2, ...a_n)$, to be denoted $\| \phi(a_1, a_2, ...a_n) \|_{\mathbf{U}}$, as follows:

(i) If ϕ is an atomic formula, then $\| a = b \|_{\mathbf{U}} = \| a = b \|$ and $\| a \in b \|_{\mathbf{U}} = \| a \in b \|$.

(ii) If ϕ is a negation, conjunction, or disjunction, then

$$\| \neg\phi(a_1, a_2, ...a_n) \|_{\mathbf{U}} = \neg \| \phi(a_1, a_2, ...a_n) \|_{\mathbf{U}},$$

$$\| \phi_1(a_1, a_2, ...a_n) \wedge \phi_2(a_1, a_2, ...a_n) \|_{\mathbf{U}} = \| \phi_1(a_1, a_2, ...a_n) \|_{\mathbf{U}} \wedge \| \phi_2(a_1, a_2, ...a_n) \|_{\mathbf{U}},$$

$$\| \phi_1(a_1, a_2, ...a_n) \vee \phi_2(a_1, a_2, ...a_n) \|_{\mathbf{U}} = \| \phi_1(a_1, a_2, ...a_n) \|_{\mathbf{U}} \vee \| \phi_2(a_1, a_2, ...a_n) \|_{\mathbf{U}},$$

$$\| \bigwedge_{\alpha<\beta} \phi_\alpha(a_1, a_2, ...a_n) \|_{\mathbf{U}} = \bigwedge_{\alpha<\beta} \| \phi_\alpha(a_1, a_2, ...a_n) \|_{\mathbf{U}}, \text{ and}$$

$$\| \bigvee_{\alpha<\beta} \phi_\alpha(a_1, a_2, ...a_n) \|_{\mathbf{U}} = \bigvee_{\alpha<\beta} \| \phi_\alpha(a_1, a_2, ...a_n) \|_{\mathbf{U}} .$$

(iii) If ϕ is $(\exists x)\psi$ or $(\forall x)\psi$, then

$$\| \exists x\phi(x, a_1, a_2, ...a_n) \|_{\mathbf{U}} = \bigvee_{a \in U} \| \phi(a, a_1, a_2, ...a_n) \|_{\mathbf{U}}, \text{ and}$$

$$\| \forall x\phi(x, a_1, a_2, ...a_n) \|_{\mathbf{U}} = \bigwedge_{a \in U} \| \phi(a, a_1, a_2, ...a_n) \|_{\mathbf{U}} .$$

Using this definition, we may assign a Boolean value to *any* sentence in the language of set theory. This Boolean value is intended to be a generalization of the ordinary notion of a truth value.

We now define the Boolean valued structure V^B that will be of interest to us.

Definition: Let B be an arbitrary complete Boolean Algebra. Then let $V^B = \bigcup_\alpha V_\alpha^B$, where α ranges over all ordinals, and V_α^B is defined inductively as follows:

(i) $V_0^B = \emptyset$

(ii) $V_{\alpha+1}^B =$ the set of all functions whose domain is a (not necessarily proper) subset of V_α^B, and whose range is a (not necessarily proper) subset of B.

(iii) $V_\alpha^B = \bigcup_{\beta<\alpha} V_\beta^B$ if α is a limit ordinal.

This construction is intended to parallel the construction of the usual V_α in 2-valued set theory, where $V_0 = \emptyset$, $V_{\alpha+1} =$ the power set of V_α, and for limit ordinals α, $V_\alpha = \bigcup_{\beta<\alpha} V_\beta$.

We also define by induction Boolean valued functions $\| x = y \|$ and $\| x \in y \|$ from $V^B \times V^B$ to B. We do this by assuming that these functions have been defined for all $x, y \in \bigcup_{\beta<\alpha} V_\beta^B$, and we show how to define them for $x, y \in V_\alpha^B$. To help us do this, we also define an auxilliary function $\| x \subseteq y \|$:

Definition:

(i) $\| \emptyset = \emptyset \| = 1$, $\| \emptyset \subseteq \emptyset \| = 1$ and $\| \emptyset \in \emptyset \| = 0$

(ii) $\| x \in y \| = \bigvee_{t \in dom(y)} (\| x = t \| \wedge y(t))$

(iii) $\| x \subseteq y \| = \bigwedge_{t \in dom(x)} (\neg x(t) \vee \| t \in y \|)$

(iv) $\| x = y \| = \| x \subseteq y \| \wedge \| y \subseteq x \|$.

Let $\mathbf{V^B}$ be the structure consisting of the set V^B and the Boolean valued functions $\| x = y \|$ and $\| x \in y \|$ from $V^B \times V^B$ to B just defined. One can then prove:

Theorem: For any complete Boolean algebra B, the following are true:

(i) $\mathbf{V^B}$ is a Boolean valued structure,

(ii) for each axiom ϕ of ZFC, $\| \phi \|_{\mathbf{V^B}} = 1$,

(iii) if $\{\phi_\alpha\}_{\alpha<\beta}$ is a set of sentences (possibly with infinitary connectives) such that $\forall \alpha < \beta$, $\| \phi_\alpha \|_{\mathbf{V^B}} = 1$, and there is a proof in the infinitary logic outlined in §2 (i.e., in $\mathcal{L}_{\infty,\omega}$) of some sentence ϕ from the premises $\{\phi_\alpha\}_{\alpha<\beta}$, then $\| \phi \|_{\mathbf{V^B}} = 1$,

(iv) for all sentences ϕ, $Con(ZFC) \rightarrow (\| \phi \wedge \neg\phi \|_{\mathbf{V^B}} = 0)$.

Fact (iii) is not usually stated in the textbooks for the case of infinitary connectives, but its proof is a straightforward generalization of the corresponding result for finitary connectives.

With this theorem, we can now prove consistency results. Imagine that ϕ is a sentence and that $\| \phi \|_{\mathbf{V^B}} = 1$. We can then immediately conclude that

$$Con(ZFC) \rightarrow Con(ZFC + \phi),$$

for if a contradiction $\psi \wedge \neg\psi$ could be deduced from $ZFC + \phi$, then we would have $\| \psi \wedge \neg\psi \|_{\mathbf{V^B}} = 1$, which would contradict fact (iv). In fact, one can use the methods developed thus far to show quite explicitly how, given a proof of a contradiction from $ZFC + \phi$, one can construct a proof of a contradiction from ZFC. (I will not present the details of this particular point.)

What we must therefore do is try to figure our which sentences ϕ satisfy $\| \phi \|_{\mathbf{V^B}} = 1$. In order to do this, it will be useful to figure out how to 'talk about' elements of the 2-valued structure V within the structure $\mathbf{V^B}$. In order to do this, we associate with every element $x \in V$ a *name* $\check{x}$ for that element in V^B as follows:

Definition: We inductively define, for each element x of V, an element $\check{x}$ of V^B as follows:

(i) $\check{\emptyset} = \emptyset$

(ii) $\check{x}$ is a function whose domain is $\{\check{y} : y \in x\}$, and is such that, for all $y \in x$, $\check{x}(\check{y}) = 1$.

The following lemma shows in what sense these names can be thought of as 'referring to' elements of V:

Lemma: The following are true:

(i) For $x, y \in V$, $\| \check{x} = \check{y} \|_{\mathbf{V^B}} = 1$ if $x = y$, and $\| \check{x} = \check{y} \|_{\mathbf{V^B}} = 0$ if $x \neq y$,

(ii) For $x, y \in V$, $\| \check{x} \in \check{y} \|_{\mathbf{V^B}} = 1$ if $x \in y$, and $\| \check{x} \in \check{y} \|_{\mathbf{V^B}} = 0$ if $x \notin y$,

(iii) For $x \in V, y \in V^B$, if $\| x = \check{y} \|_{\mathbf{V^B}} = 1$, then $x = \check{y}$,

(iv) For $x \in V, y \in V^B$, if $\| x \in \check{y} \|_{\mathbf{V^B}} = 1$, then $x = \check{z}$ for some $z \in y$. Equivalently,

$$\| (x \in \check{y}) \rightarrow \bigvee_{z \in y} (x = \check{z}) \|_{\mathbf{V^B}} = 1.$$

Thus, we may think of the mapping $i : V \rightarrow V^B$ given by $i(x) = \check{x}$ as an embedding of V into V^B.

Before we can state the next result, we need some definitions. For $x \in V$, let $T(x)$ be the sentence

$$(y \in x) \rightarrow \bigvee_{z \in x} (y = z).$$

Let $S(x)$ be the smallest set of sentences such that $T(x) \in S(x)$, and satisfying

$$\forall y, z[((T(y) \in S(x)) \wedge (z \in y)) \rightarrow (T(z) \in S(x))].$$

Let $U(x) = \bigwedge_{\phi \in S(x)} \phi$. The sentence $U(x)$ uniquely specifies the set x. Let $\bar{U}(\check{x})$ be the sentence $U(x)$ with all x, y, z replaced by $\check{x}, \check{y}, \check{z}$. The following is easily proved:

Lemma: For all $x \in V$, $\| \bar{U}(\check{x}) \|_{\mathbf{V^B}} = 1$.

Using this, we prove the following key lemma:

Lemma: Let ϕ be a sentence in the language of set theory, and let $a_1, a_2, \ldots a_n \in V$.

If $\| \phi(\check{a}_1, \check{a}_2, \ldots \check{a}_n) \|_{\mathbf{V^B}} = 1$, then

$$Con(ZFC) \rightarrow Con(ZFC + \phi(a_1, a_2, ..., a_n) + \bigwedge_{i \leq n} U(a_i)).$$

Note that by $Con(T)$, we mean consistency in the sense outlined in §2, i.e., relative to the axioms as rules of inference of the infinitary logic $\mathcal{L}_{\infty,\omega}$, and not just first-order logic. (The analog of this lemma for the finitary case is all one finds presented in the usual textbooks.) Note also that the purpose of the $\bigwedge_{i\leq n} U(a_i)$ is to fully specify the meaning of the $a_1, a_2, ...a_n$, which would otherwise be uninterpreted terms.

The next lemma is also of great importance:

Lemma: Let B be a complete Boolean Algebra. Let $G \in V^B$ be defined as the function with domain $\{\check{u} : u \in B\}$, such that $G(\check{u}) = u$ for every $u \in B$. Finally, let D be the power set of B in V, i.e., $D = \mathcal{P}^V(B)$. Then:

$$\| G \text{ is a } \check{D} \text{ generic ultrafilter on } \check{B} \|_{\mathbf{V^B}} = 1.$$

Putting all of this together, we get the following:

Theorem: $Con(ZFC) \rightarrow Con(ZFC +$ there is a $\mathcal{P}^V(B)$ generic ultrafilter on $B + U(B)$.)

This is the central result of the technique of forcing. Note that it applies for any complete Boolean Algebra B. Given a sentence ϕ for which we want to prove

$$Con(ZFC) \rightarrow Con(ZFC + \phi),$$

it therefore suffices to find a complete Boolean Algebra B such that

$$(ZFC + \text{there is a } \mathcal{P}^V(B) \text{ generic ultrafilter on } B + U(B)) \rightarrow \phi.$$

This is the strategy of all independence proofs obtained by the method of forcing.

In our case, ϕ is the sentence 'there is a Dedekind cut of $\mathbb{Q}^V$ not contained in $\mathbb{R}^V$.' We now construct a Boolean Algebra B for which

$$(ZFC + \text{there is a } \mathcal{P}^V(B) \text{ generic ultrafilter on } B + U(B)) \rightarrow \phi.$$

The construction is a standard one. Let P be the partially ordered set of finite strings of 0s and 1s, ordered under reverse inclusion, i.e., $\sigma_1 < \sigma_2$ iff σ_1 properly extends σ_2. (Thus, the null string $\emptyset$ is greater than every element in P.) So P is an infinite, upside-down, binary tree.

If $U \subseteq P$, say that U is a *cut* if $\sigma \in U$ and $\sigma' < \sigma$ implies $\sigma' \in U$. Say that a cut U is *regular* just in case U is a cut, and

$$\forall x[(x \notin U) \to (\exists y < x)(\forall z \leq y)(z \notin U)].$$

We then have the following result:

Lemma: For all cuts U, there is a unique smallest cut $\bar{U}$ such that $U \subseteq \bar{U}$ and $\bar{U}$ is regular.

Let B be the set of all regular cuts U of P. For $U_1, U_2 \in B$, define $U_1 \wedge U_2 = U_1 \cap U_2$, $U_1 \vee U_2 = \overline{U_1 \cup U_2}$, and $\neg U = \{\sigma : \forall \sigma' \leq \sigma(\sigma' \notin U)\}$. Then we have the following:

Lemma: Under the operations $\wedge$, $\vee$, and $\neg$ just defined, B is a complete Boolean Algebra.

For $\sigma \in P$, let $U_\sigma = \{\sigma' : \sigma' \leq \sigma\}$. Then U_σ is a regular cut. For each $n \in \mathbb{N}$, let $W^n = \{U_\sigma : \text{ length}(\sigma) \geq n\}$. Then we can prove the following:

Lemma: If G is a $\mathbb{P}^V(B)$ generic ultrafilter on B, then for each $n \in \mathbb{N}$, $G \cap W^n \neq \emptyset$.

What this means is that if G is a $\mathbb{P}^V(B)$ generic ultrafilter on B, then for each $n \in \mathbb{N}$ there is a $\sigma_n \in P$ such that $U_{\sigma_n} \in G$ and $\text{length}(\sigma_n) \geq n$. In fact, we have the following:

Lemma: Let $n, n' \in \mathbb{N}$. If $\text{length}(\sigma_n) \leq \text{length}(\sigma'_n)$, then σ'_n extends σ_n, i.e., $\sigma'_n \leq \sigma_n$.

What this means is that we can form a (countably) infinite string σ of 0s and 1s such that, for all n, $U_{\sigma|n} \in G$, where $\sigma|n$ is just the substring of σ consisting of its first n digits.

Let π be any (countably) infinite string of 0s and 1s. Define:

$$F^\pi = \{U : U \subseteq P, U \text{ is a regular cut, and } \exists n(U_{\pi|n} \subseteq U)\}.$$

We then have:

Lemma: If G is a $\mathbb{P}^V(B)$ generic ultrafilter on B, and $\pi \in V$ is a (countably) infinite string of 0s and 1s, then $F^\pi \not\subseteq G$.

Proof. It suffices to show that $\exists U \in F^\pi (U \notin G)$. Assume to the contrary that $\forall U \in F^\pi (U \in G)$. Then $\forall n (U_{\pi|n} \in G)$. But if $\pi \in V$, then $\{U_{\pi|n}\} \in \mathcal{P}^V(B)$, and so we must have $\bigwedge_n U_{\pi|n} \in G$. But $\bigwedge_n U_{\pi|n} = 0$, and so $0 \in G$, contradiction. $\square$

From the fact that $\exists \sigma (F^\sigma \subseteq G)$ and $\forall \pi \in V (F^\pi \nsubseteq G)$, we can conclude that there exists a countably infinite string of 0s and 1s such that $\sigma \notin V$. We can think of σ as a real number that is not contained in $\mathbb{R}^V$. Thus we have proven:

Theorem: For the complete Boolean Algebra B just described,

$$(ZFC + \text{ there is a } \mathcal{P}^V(B) \text{ generic ultrafilter on } B + U(B)) \to \phi,$$

where $\phi =$ there is a Dedekind cut of $\mathbb{Q}^V$ not contained in $\mathbb{R}^V$.

We therefore have our Main Theorem:

Main Theorem: $Con(ZFC) \to Con(ZFC+$there is a Dedekind cut of $\mathbb{Q}^V$ not contained in $\mathbb{R}^V$.)

BIBLIOGRAPHY

[1] Almog., J [1991], The Plentitude of Structures and the Scarcity of Possibilities, *The Journal Of Philosophy*, 88(11), 620–622.

[2] Aristotle., [1988], *Metaphysics*, Clarendon.

[3] Atiyah, M. et al. [1994], 'Responses to 'Theoretical Mathematics: Toward a Cultural Synthesis of Mathematics and Theoretical Physics', by A. Jaffe and F. Quinn', *Bulletin of the American Mathematical Society*, 30, pp. 178-207.

[4] Auden, W., and Kronenberger, L., (eds.) [1996], *The Viking Book of Aphorisms*, New York: Viking Press.

[5] Azzouni, J., [2000], 'Applying Mathematics: An Attempt to Design a Philosophical Problem', *Monist*, **83**(2).

[6] Barr, W., [1971], 'A Syntactic and Semantic Analysis of Idealizations in Science', *Philosophy of Science*, 38, 258–272.

[7] Barwise. J., [1975], *Admissible Sets and Structures*, Springer-Verlag, Berlin.

[8] Bell, E., [1937], *Men of Mathematics*, New York.

[9] Beller, M. [2001], *Quantum Dialogue : The Making of a Revolution*, Chicago: University of Chicago Press.

[10] Bergson, H., [1903], 'Introduction to Metaphysics', reprinted in H. Bergson, [1946], *The Creative Mind*, Citadel Press.

[11] Bridgman, P., [1951], 'The Nature of Some of Our Physical Concepts', *The British Journal for the Philosophy of Science*, 1 (4), pp 257-272.

[12] Bridgman, P. [1959], 'How Much Rigor is Possible in Physics?', in L. Henkin, P. Suppes, and A. Tarski (eds), 1959, *The Axiomatic Method, with Special Reference to Geometry and Physics,* Amsterdam: North-Holland Publishing Company.

[13] Brown, B. [1990], 'How to be Realistic about Inconsistency in Science', *Studies in the History and Philosophy of Science*, vol 21, no 2, pp. 281-44.

[14] Callender, C., [2001], Review of *Explaining Chaos* by P. Smith., Mind 110 (439) 839–844.

[15] Cameron, R. H. [1960], 'A Family of Integrals Serving to Connect the Wiener and Feynman Integrals', *Journal of Mathematics and Physics*, 39, pp. 126-140.

[16] Carnap, R. [1967], *The Logical Structure of the World*, Berkeley: University of California Press.

[17] Cartwright, N., [1983], *How the Laws of Physics Lie*, Oxford University Press.

[18] Caspar, M., ed. [1937], *Johannes Kepler Gesammelte Werke*, Munich.

[19] Chang, C., and Kiesler, J., [1990], *Model Theory*, Elsevier Science.

[20] Cohen, P., [1963], The Independence of the Continuum Hypothesis I, *Proceedings of the National Academy of Science, USA*, **50**, 1143–1148.

[21] Cohen, P., [1964], The Independence of the Continuum Hypothesis II, *Proceedings of the National Academy of Science, USA*, **51**, 105–110.

[22] Coleman, A., [1997], 'Groups and Physics - Dogmatic Opinions of a Senior Citizen', *Notices of the American Mathematical Society*, **44**, pp. 8-17.

[23] Dennett, D. [1987], *The Intentional Stance*, MIT Press.

[24] Dicke, R. and Wittke, J. [1960], *Introduction to Quantum Mechanics*, Reading: Addison-Wesley Publishing Company.

[25] Dickman, M., [1975], *Large Infinitary Languages*, North-Holland, Amsterdam.

[26] Dirac, P. [1958], *The Principles of Quantum Mechanics*, New York: Oxford University Press.

[27] Duhem, P. [1977], *The Aim and Structure of Physical Theory*, Princeton: Princeton University Press.

[28] Dyson, F., [1964], 'Mathematics in the Physical Sciences', *Scientific American*, **211**, pp. 128-146.

[29] Earman, J., [1970], 'Who's Afraid of Absolute Space?', *Australasian Journal of Philosophy*, 48, 287–319.

[30] Einstein, A., [1933], 'On the Methods of Theoretical Physics', in Einstein, A., [1954], *Ideas and Opinions*, New York, Bonanza, pp. 270-276.

[31] Farmer, J., Ott. E., and Yorke, J., [1983], 'The Dimension of Strange Attractors', *Physica D*, (7), pp 153-180.

[32] Feynman, R. and Hibbs, A. [1965], *Quantum Mechanics and Path Integrals*, New York: McGraw-Hill, Inc.

[33] Feynman, R. et al. [1985], *The Feynman Lectures on Physics*, Reading: Addison Wesley Longman, Inc.

[34] Feynman, R., [1967], *The Character of Physical Law*, MIT Press.

[35] French, S., [2000], 'The Reasonable Effectiveness of Mathematics: Partial Structures and the Application of Group Theory to Physics', *Synthese*, **125**, pp. 103-120.

[36] Gardner, M. [1985], *Wheels, Life and Other Mathematical Amusements.*, W. H. Freeman and Co.

[37] Gelfand, I. and Yaglom, A., [1956], 'Integration in Function Space and Its Application in Quantum Physics', *Uspekhi Mat. Nauk.*, Vol 11, pp. 77-114.

[38] Geroch. R., [20xx] Path Integrals, Unpublished Manuscript.

[39] Glimm, J. and Jaffe, A. [1981], *Quantum Mechanics - A Functional Integral Point of View*, Heidelberg: Springer Verlag.

[40] Gordon, I., and Sorkin, S., (eds.) [1959], *The Armchair Science Reader*, New York, Simon and Schuster.

[41] Hamming, R., [1980] 'The Unreasonable Effectiveness of Mathematics', *American Mathematics Monthly*, **87**, pp. 81-90.

[42] Heaviside, O. [1971], *Electromagnetic Theory, Vol II*, New York, Chelsea Publishing Company.

[43] Hempel, C. [1965], *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, New York: The Free Press.

[44] Hempel, C., [1965], 'A Logical Appraisal of Operationism', reprinted as Chapter 5 of *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, The Free Press.

[45] Hersh, R., [1990] 'Inner Vision, Outer Truth', in R. Mickens (ed.) *Mathematics and Science*, Singapore: World Scientific Press, pp. 64-72.

[46] Jaffe, A. and Quinn, F., (1993) 'Theoretical Mathematics: Toward a Cultural Synthesis of Mathematics and Theoretical Physics', *Bulletin of the American Mathematical Society*, **29**, pp. 1-13.

[47] Jammer, M. [1966], *The Conceptual Development of Quantum Mechanics*, New York: McGraw Hill Book Company.

[48] Jech, T., [1997], *Set Theory*, Springer-Verlag.

[49] Johnson, G. and Lapidus, M., [2002], *The Feynman Integral and Feynman's Operational Calculus*, Oxford University Press.

[50] Jouvenel, B., [1967], *The Art of Conjecture*, New York, Basic Books.

[51] Kijowski, J., [1995], 'Can Classical Electrodynamics Be Renormalized?', *Reports on Mathematical Physics*, 35, 207-223.

[52] Kitcher, P. [1981], 'Mathematical Rigor – Who needs it?', *Nous*, vol 15, no 4, pp. 469-493.

[53] Kunen, K., [1980], *Set Theory*, Elsevier Science Publishing Co.

[54] Landau, L. and Lifshitz, E. [1938]: *Statistical Physics*, Oxford: Oxford University Press.

[55] Machamer, P., (ed.) [1998], *The Cambridge Companion to Galileo*, Cambridge.

[56] Mac Lane, S., [1990], 'The Reasonable Effectiveness of Mathematics', in R. Mickens (ed.) *Mathematics and Science*, Singapore: World Scientific Press, pp. 115-135.

[57] Malament, D. [1995], 'Is Newtonian Cosmology Really Inconsistent?', *Philosophy of Science*, 62, pp. 489-510.

[58] Maxwell, J.C., [1873], *A Treatise on Electricity and Magnetism*, Clarendon Press.

[59] Maxwell, J. C., [1995], 'On Action at a Distance', in P. Harman (ed.) *The Scientific Letters and Papers of James Clerk Maxwell*, Cambridge University Press.

[60] Maxwell, J. C., [1952], 'On physical lines of force', in Niven, W., *The scientific papers of James Clerk Maxwell*, Cambridge University Press, v. 1, 488-489.

[61] Mehra, J., [1996], *The Beat of a Different Drum: The Life and Science of Richard Feynman*, Oxford University Press.

[62] Montgomery, R., [2001], 'A New Solution to the Three Body Problem', *Notices of the American Mathematical Society*, 48(May): 471.

[63] Moschovakis Y., [1980], *Descriptive Set Theory*, North Holland Publishing Company.

[64] Nagel, E., [1979], 'Impossible Numbers', in *Teleology Revisited*, New York, Columbia.

[65] Nelson E., [1964], Feynman Integrals and the Schrödinger Equation, *J. Math. Phys.*, 5 , 332-343.

[66] Nietzsche, F., [1990], *Beyond Good and Evil*, Dover.

[67] Norton, J., [2003], 'A Paradox in Newtonian Gravitation Theory II', forthcoming.

[68] Nowak, L., [1972], Laws of Science, Theories, Measurement *Philosophy of Science*, 39, 533–548.

[69] Papineau, D., [1979], *Theory and Meaning*, Oxford University Press.

[70] Peirce, C., [1958] *Collected Papers of Charles Sanders Pierce*, ed. A. Burks. Vol. 7., Harvard University Press.

[71] Polkinghorne, J., [1990] 'The Reason Within and the Reason Without', in R. Mickens (ed.) *Mathematics and Science*, Singapore: World Scientific Press, pp. 173-182.

[72] Popper, K. [1992], *The Logic of Scientific Discovery*, New York: Routledge.

[73] Priest, G., [2000], 'Could everything be True?', *Australasian Journal of Philosophy*, 78, 189-95.

[74] Putnam, H., [1981], *Reason, Truth and History.* Cambridge University Press.

[75] Rassias, G., [1991], *The mathematical heritage of C. F. Gauss*, Singapore.

[76] Redhead, M., [1980], 'Models in Physics', *British Journal of the Philosophy of Science*, 31, 145–163.

[77] Rivers, R. J. [1987], *Path Integral Methods in Quantum Field Theory*, Cambridge: Cambridge University Press.

[78] Rosen, R., [1990] 'The Modelling Relation and Natural Law', in R. Mickens (ed.) *Mathematics and Science*, Singapore: World Scientific Press, pp. 183-199.

[79] Rudin, W., [1970], *Real and Complex Analysis*, McGraw-Hill.

[80] Rudin, W., [1964], *Principles of Mathematical Analysis*, McGraw-Hill.

[81] Russell, B., [1913], 'On the Notion of Cause', *Proceedings of the Aristotelian Society*, 13, pp. 1-26.

[82] Scheibe, E., [1994], 'On the Mathematical Overdetermination of Physics', in E. Rudolph (ed.) *Philosophy, Mathematics and Modern Physics: A Dialogue*, Springer Verlag.

[83] Schwartz, J., [1962], 'The Pernicious Influence of Mathematics on Science', in *Logic, Methodology and Philosophy of Science, Proceedings of the 1960 International Congress*, Stanford University Press, pp. 356-360.

[84] Schwartz, L. [1945], 'Generalization de la notion de fonction, de derivation, de transformation de Fourier et applications mathematicques et physiques', *Annales de l'Universite de Grenoble*, 21, pp. 57-74.

[85] Scott, D., [1965], 'Logic with denumerably long formulas and finite strings of quantifiers', in *The Theory of Models*, Addison, J., Henkin, L., and Tarki, A. (eds), North-Holland, Amsterdam, pp 324-341.

[86] Shapiro, S., [1983], 'Mathematics and Reality', *Philosophy of Science*, **50**, pp. 523-548.

[87] Shoenfield, J., [2001], *Mathematical Logic*, AK Peters.

[88] Smith, J. [1998], 'Inconsistency and Scientific Reasoning', *Studies in the History and Philosophy of Science*, vol 19, no 4, pp. 429-445.

[89] Smith, P., [1998], *Explaining Chaos*, Cambridge University Press.

[90] Steiner, M., [1992], 'Mathematical Rigor in Physics', in M. Detlefsen (ed.), *Proof and Knowledge in Mathematics*, Routledge.

[91] Steiner, M. [1989], 'The Application of Mathematics to Natural Science', *Journal of Philosophy*, 86, pp. 449-480.

[92] Steiner, M., [1998] *The Applicability of Mathematics as a Philosophical Problem*, Harvard University Press.

[93] Swanson, M., [1992], *Path Integrals and Quantum Processes*, Academic Press.

[94] Ulam, S., [1969], 'The Applicability of Mathematics', in *The Mathematical Sciences, ed. Committee on Support of Research in the Mathematical Sciences (COSRIMS) of the National Research Council*, Cambridge, Mass, MIT Press, pp. 1-6.

[95] von Neumann, J. [1996], *Mathematical Foundations of Quantum Mechanics*, Princeton: Princeton University Press.

[96] Wang, Q., and Young, L., [2002], 'From Invariant Curves to Strange Attractors', *Communications in Mathematical Physics*, 225 (2), 275–304.

[97] Wang, Q., and Young, L., [2001], 'Strange Attractors with One Direction of Instability ', *Communications in Mathematical Physics*, 218 (1), 1–97.

[98] Weinberg, S., [1986], 'Lecture on the Applicability of Mathematics', *Notices of the American Mathematical Society*, **33**(5), pp. 725-728.

[99] Weston, T., [1992], 'Approximate Truth and Scientific Realism', *Philosophy of Science*, 59, 53–74.

[100] Weston, T., [1987], 'Approximate Truth', *Journal of Philosophical Logic*, 16, 203–227.

[101] Wigner, E., [1960], 'The Unreasonable Effectiveness of Mathematics in the Natural Sciences', *Communications on Pure and Applied Mathematics*, **13**, pp. 1-14.

[102] Wilson, M., [2000] 'On the Mathematics of Spilt Milk', in E. Grosholz, H. Breger (eds.), *The Growth of Mathematical Knowledge*, 143-152. Kluwer, pp. 143-151.

[103] Wilson, M., [2000] 'The Unreasonable Uncooperativeness of Mathematics in the Natural Sciences', *The Monist*, **83**(2), pp 296-314.

[104] Wussing, H., and Arnold, W., [1983], *Biographien bedeutender Mathematiker*, Berlin.

[105] Zee, A., [1990], 'The Effectiveness of Mathematics in Fundamental Physics', in R. Mickens (ed.) *Mathematics and Science*, Singapore: World Scientific Press, pp. 307-323.